

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
21 March 2002 (21.03.2002)

PCT

(10) International Publication Number  
**WO 02/21996 A2**

- (51) International Patent Classification<sup>7</sup>: **A61B**
- (21) International Application Number: PCT/US01/28716
- (22) International Filing Date:  
14 September 2001 (14.09.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
09/663,733 15 September 2000 (15.09.2000) US  
09/930,020 14 August 2001 (14.08.2001) US
- (71) Applicant (*for all designated States except US*): **EOS BIOTECHNOLOGY, INC.** [US/US]; 225 A Gateway Boulevard, South San Francisco, CA 94080 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **GISH, Kurt, C.** [US/US]; 40 Perego Terrace #2, San Francisco, CA 94131 (US). **MACK, David, H.** [US/US]; 2076 Monterey Avenue, Menlo Park, CA 94025 (US). **WILSON, Keith, E.** [US/US]; 219 Jeter Street, Redwood City, CA 94062 (US).
- (74) Agents: **BASTIAN, Kevin, L.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, Eighth Floor, San Francisco, CA 94111 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**  
— *without international search report and to be republished upon receipt of that report*
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: METHODS OF DIAGNOSIS OF COLORECTAL CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR COLORECTAL CANCER MODULATORS

(57) Abstract: Described herein are methods that can be used for diagnosis and prognosis of colorectal cancer. Also described herein are methods that can be used to screen candidate bioactive agents for the ability to modulate colorectal cancer. Additionally, methods and molecular targets (genes and their products) for therapeutic intervention in colorectal and other cancers are described.



**WO 02/21996 A2**

## Methods of Diagnosis of Colorectal Cancer, Compositions and Methods of Screening for Colorectal Cancer Modulators

### CROSS-REFERENCES TO RELATED APPLICATIONS

[01] This application is a continuation in part of US Patent Application USSN 09/663,733 filed September 15, 2000, and US Patent Application filed August 14, 2001 USSN, not yet known, which are incorporated herein by reference in their entirety.

### FIELD OF THE INVENTION

[02] The invention relates to the identification of expression profiles and the nucleic acids involved in colorectal cancer, and to the use of such expression profiles and nucleic acids in diagnosis and prognosis of colorectal cancer. The invention further relates to methods for identifying and using candidate agents and/or targets which modulate colorectal cancer.

### BACKGROUND OF THE INVENTION

[03] Cancer of the colon and/or rectum (referred to as "colorectal cancer") are significant in Western populations and particularly in the United States. Cancers of the colon and rectum occur in both men and women most commonly after the age of 50. These develop as the result of a pathologic transformation of normal colon epithelium to an invasive cancer. There have been a number of recently characterized genetic alterations that have been implicated in colorectal cancer, including mutations in two classes of genes, tumor-suppressor genes and proto-oncogenes, with recent work suggesting that mutations in DNA repair genes may also be involved in tumorigenesis. For example, inactivating mutations of both alleles of the adenomatous polyposis coli (APC) gene, a tumor suppressor gene, appears to be one of the earliest events in colorectal cancer, and may even be the initiating event. Other genes implicated in colorectal cancer include the MCC gene, the p53 gene, the DCC (deleted in colorectal carcinoma) gene and other chromosome 18q genes, and genes in the TGF- $\beta$  signaling pathway. For a review, see *Molecular Biology of Colorectal Cancer*, pp. 238-299, in *Curr. Probl. Cancer*, Sept/Oct 1997; see also Willams, *Colorectal Cancer* (1996); Kinsella & Schofield, *Colorectal Cancer: A Scientific Perspective* (1993); *Colorectal*

*Cancer: Molecular Mechanisms, Premalignant State and its Prevention* (Schmiegel & Scholmerich eds., 2000); *Colorectal Cancer: New Aspects of Molecular Biology and Their Clinical Applications* (Hanski *et al.*, eds 2000); McArdle *et al.*, *Colorectal Cancer* (2000); Wanebo, *Colorectal Cancer* (1993); Levin, *The American Cancer Society: Colorectal Cancer* (1999); *Treatment of Hepatic Metastases of Colorectal Cancer* (Nordlinger & Jaeck eds., 1993); *Management of Colorectal Cancer* (Dunitz *et al.*, eds. 1998); *Cancer: Principles and Practice of Oncology* (Devita *et al.*, eds. 2001); *Surgical Oncology: Contemporary Principles and Practice* (Kirby *et al.*, eds. 2001); Offit, *Clinical Cancer Genetics: Risk Counseling and Management* (1997); *Radioimmunotherapy of Cancer* (Abrams & Fritzberg eds. 2000); Fleming, *AJCC Cancer Staging Handbook* (1998); *Textbook of Radiation Oncology* (Leibel & Phillips eds. 2000); and *Clinical Oncology* (Abeloff *et al.*, eds. 2000).

[04] Imaging of colorectal cancer for diagnosis has been problematic and limited. In addition, metastasis of the tumor to the lumen, and metastasis of tumor cells to regional lymph nodes are important prognostic factors (*see, e.g., PET in Oncology: Basics and Clinical Application* (Ruhlmann *et al.* eds. 1999). For example, five year survival rates drop from 80 percent in patients with no lymph node metastases to 45 to 50 percent in those patients who do have lymph node metastases. A recent report showed that micrometastases can be detected from lymph nodes using reverse transcriptase-PCR methods based on the presence of mRNA for carcinoembryonic antigen, which has previously been shown to be present in the vast majority of colorectal cancers but not in normal tissues. Liefers *et al.*, *New England J. of Med.* 339(4):223 (1998).

[05] Thus, methods that can be used for diagnosis and prognosis of colorectal cancer would be desirable. Accordingly, provided herein are methods that can be used in diagnosis and prognosis of colorectal cancer. Further provided are methods that can be used to screen candidate bioactive agents for the ability to modulate colorectal cancer. Additionally, provided herein are molecular targets for therapeutic intervention in colorectal and other cancers.

#### BRIEF SUMMARY OF THE INVENTION

[06] The present invention provides novel methods for diagnosis and prognosis evaluation for colorectal cancer, as well as methods for screening for compositions which modulate colorectal cancer. Methods of treatment of colorectal cancer, as well as compositions, are also provided herein.

[07] In one aspect, a method of screening drug candidates comprises providing a cell that expresses an expression profile gene selected from those of Table I. The method further includes adding a drug candidate to the cell and determining the effect of the drug candidate on the expression of the expression profile gene.

5 [08] In one embodiment, the method of screening drug candidates includes comparing the level of expression in the absence of the drug candidate to the level of expression in the presence of the drug candidate, wherein the concentration of the drug candidate can vary when present, and wherein the comparison can occur after addition or removal of the drug candidate. In a preferred embodiment, the cell expresses at least two  
10 expression profile genes. The profile genes may show an increase or decrease.

[09] Also provided herein is a method of screening for a bioactive agent capable of binding to a colorectal cancer modulator protein, the method comprising combining the colorectal cancer modulator protein and a candidate bioactive agent, and determining the binding of the candidate agent to the colorectal cancer modulator protein.

15 Preferably the colorectal cancer modulator protein is a product encoded by a gene of Table 1 or Table 2.

[10] Further provided herein is a method for screening for a bioactive agent capable of modulating the activity of a colorectal cancer modulator protein. In one embodiment, the method comprises combining the colorectal cancer modulator protein and a  
20 candidate bioactive agent, and determining the effect of the candidate agent on the bioactivity of the colorectal cancer modulator protein. Preferably the colorectal cancer modulator protein is a product encoded by a gene of Table 1 or Table 2.

[11] Also provided is a method of evaluating the effect of a candidate colorectal cancer drug comprising administering the drug to a transgenic animal expressing or  
25 over-expressing the colorectal cancer modulator protein, or an animal lacking the colorectal cancer modulator protein, for example as a result of a gene knockout.

[12] Additionally, provided herein is a method of evaluating the effect of a candidate colorectal cancer drug comprising administering the drug to a patient and removing a cell sample from the patient. The expression profile of the cell is then determined. This  
30 method may further comprise comparing the expression profile to an expression profile of a healthy individual. In a preferred embodiment, said expression profile includes a gene of Table 1 or Table 2.

[13] Moreover, provided herein is a biochip comprising one or more nucleic acid segments of Table 1 or Table 2, wherein the biochip comprises fewer than 1000 nucleic acid probes. Preferable at least two nucleic acid segments are included.

[14] Furthermore, a method of diagnosing a disorder associated with colorectal cancer is provided. The method comprises determining the expression of a gene of Table 1 or Table 2, in a first tissue type of a first individual, and comparing the distribution to the expression of the gene from a second normal tissue type from the first individual or a second unaffected individual. A difference in the expression indicates that the first individual has a disorder associated with colorectal cancer.

[15] In another aspect, the present invention provides an antibody which specifically binds to a protein encoded by a nucleic acid of Table 1 or Table 2 or a fragment thereof. Preferably the antibody is a monoclonal antibody. The antibody can be a fragment of an antibody such as a single stranded antibody as further described herein, or can be conjugated to another molecule. In one embodiment, the antibody is a humanized antibody.

[16] In one embodiment a method for screening for a bioactive agent capable of interfering with the binding of a colorectal cancer modulating protein (colorectal cancer modulator protein) or a fragment thereof and an antibody which binds to said colorectal cancer modulator protein or fragment thereof. In a preferred embodiment, the method comprises combining a colorectal cancer modulator protein or fragment thereof, a candidate bioactive agent and an antibody which binds to said colorectal cancer modulator protein or fragment thereof. The method further includes determining the binding of said colorectal cancer modulator protein or fragment thereof and said antibody. Wherein there is a change in binding, an agent is identified as an interfering agent. The interfering agent can be an agonist or an antagonist. Preferably, the agent inhibits colorectal cancer.

[17] In a further aspect, a method for inhibiting colorectal cancer is provided. The method can be performed in vitro or in vivo, preferably in vivo to an individual. In a preferred embodiment the method of inhibiting colorectal cancer is provided to an individual with cancer. As described herein, methods of inhibiting colorectal cancer can be performed by administering an inhibitor of the activity of a protein encoded by a nucleic acid of Table 1 or Table 2, including an antisense molecule to the gene or its gene product.

[18] Also provided herein are methods of eliciting an immune response in an individual. In one embodiment a method provided herein comprises administering to an individual a composition comprising a colorectal cancer modulating protein, or a fragment

thereof. In another embodiment, the protein is encoded by a nucleic acid selected from those of Table 1 or Table 2. In another aspect, said composition comprises a nucleic acid comprising a sequence encoding a colorectal cancer modulating protein, or a fragment thereof.

5           [19] Further provided herein are compositions capable of eliciting an immune response in an individual. In one embodiment, a composition provided herein comprises a colorectal cancer modulating protein, preferably encoded by a nucleic acid of Table 1 or Table 2, or a fragment thereof, and a pharmaceutically acceptable carrier. In another embodiment, said composition comprises a nucleic acid comprising a sequence  
10 encoding a colorectal cancer modulating protein, preferably selected from the nucleic acids of Table 1 or Table 2 and a pharmaceutically acceptable carrier.

          [20] Also provided are methods of neutralizing the effect of a colorectal cancer protein, or a fragment thereof, comprising contacting an agent specific for said protein with said protein in an amount sufficient to effect neutralization. In another embodiment, the  
15 protein is encoded by a nucleic acid selected from those of Table 1 or Table 2.

          [21] In another aspect of the invention, a method of treating an individual for colorectal cancer is provided. In one embodiment, the method comprises administering to said individual an inhibitor of a colorectal cancer modulating protein. In another  
20 embodiment, the method comprises administering to a patient having colorectal cancer an antibody to a colorectal cancer modulating protein conjugated to a therapeutic moiety. Such a therapeutic moiety can be a cytotoxic agent or a radioisotope.

          [22] Compounds and compositions are also provided. Other aspects of the invention will become apparent to the skilled artisan by the following description of the invention.

#### 25                               BRIEF DESCRIPTION OF THE DRAWINGS

                                  [NOT APPLICABLE]

#### DETAILED DESCRIPTION OF THE INVENTION

          [23] The present invention provides novel methods for diagnosis and  
30 prognosis evaluation for colorectal cancer, as well as methods for screening for compositions which modulate colorectal cancer. The methods herein are related to those of U.S. Patent Application Serial No. 09/525,993 and International Patent Application No. PCT/US00/07044, each of which is incorporated herein in its entirety.

[24] By "colorectal cancer" herein is meant a colon and/or rectal tumor or cancer that is classified as Dukes stage A or B as well as metastatic tumors classified as Dukes stage Cor D (see, e.g., Cohen *et al.*, *Cancer of the Colon*, in *Cancer: Principles and Practice of Oncology*, pp. 1144-1197 (Devita *et al.*, eds., 5<sup>th</sup> ed. 1997); see also *Harrison's Principles of Internal Medicine*, pp. 1289-129 (Wilson *et al.*, eds., 12<sup>th</sup> ed., 1991).

"Treatment, monitoring, detection or modulation of colorectal cancer" includes treatment, monitoring, detection, or modulation of colorectal disease in those patients who have colorectal disease (Dukes stage A, B, C or D) in which gene expression from a gene in Table 1 or 2, is increased or decreased, indicating that the subject is more likely to progress to metastatic disease than a patient who does not have an increase or decrease in gene expression of a gene in Table 1 or 2. In Dukes stage A, the tumor has penetrated into, but not through, the bowel wall. In Dukes stage B, the tumor has penetrated through the bowel wall but there is not yet any lymph involvement. In Dukes stage C, the cancer involves regional lymph nodes. In Dukes stage D, there is distant metastasis, e.g., liver, lung, etc.

[25] Table 1 provides unigene cluster identification numbers for the nucleotide sequence of genes that exhibit increased expression in colorectal cancer samples. Tables 1 also provides an exemplar accession number that provides a nucleotide sequence that is part of the unigene cluster. Table 2 provides the nucleic acid and protein sequence of the CBF9 gene as well as the Unigene and Exemplar accession numbers for CBF9.

[26] In one aspect, the expression levels of genes are determined in different patient samples for which either diagnosis or prognosis information is desired, to provide expression profiles. An expression profile of a particular sample is essentially a "fingerprint" of the state of the sample; while two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is unique to the state of the cell. That is, normal tissue may be distinguished from colorectal cancer tissue, and within colorectal cancer tissue, different prognosis states (good or poor long term survival prospects, for example) may be determined. By comparing expression profiles of colon tissue in known different states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained. The identification of sequences that are differentially expressed in colorectal cancer versus normal colon tissue, as well as differential expression resulting in different prognostic outcomes, allows the use of this information in a number of ways. For example, the evaluation of a particular treatment regime may be evaluated: does a chemotherapeutic drug act to improve the long-term

prognosis in a particular patient. Similarly, diagnosis may be done or confirmed by comparing patient samples with the known expression profiles. Furthermore, these gene expression profiles (or individual genes) allow screening of drug candidates with an eye to mimicking or altering a particular expression profile; for example, screening can be done for drugs that suppress the colorectal cancer expression profile or convert a poor prognosis profile to a better prognosis profile. This may be done by making biochips comprising sets of the important colorectal cancer genes, which can then be used in these screens. These methods can also be done on the protein basis; that is, protein expression levels of the colorectal cancer proteins can be evaluated for diagnostic and prognostic purposes or to screen candidate agents. In addition, the colorectal cancer nucleic acid sequences can be administered for gene therapy purposes, including the administration of antisense nucleic acids, or the colorectal cancer proteins (including antibodies and other modulators thereof) administered as therapeutic drugs.

[27] Thus the present invention provides nucleic acid and protein sequences that are differentially expressed in colorectal cancer, herein termed "colorectal cancer sequences". As outlined below, colorectal cancer sequences include those that are up-regulated (i.e. expressed at a higher level) in colorectal cancer, as well as those that are down-regulated (i.e. expressed at a lower level) in colorectal cancer. In a preferred embodiment, the colorectal cancer sequences are from humans; however, as will be appreciated by those in the art, colorectal cancer sequences from other organisms may be useful in animal models of disease and drug evaluation; thus, other colorectal cancer sequences are provided, from vertebrates, including mammals, including rodents (rats, mice, hamsters, guinea pigs, etc.), primates, farm animals (including sheep, goats, pigs, cows, horses, etc). colorectal cancer sequences from other organisms may be obtained using the techniques outlined below.

[28] Colorectal cancer sequences can include both nucleic acid and amino acid sequences. In a preferred embodiment, the colorectal cancer sequences are recombinant nucleic acids. By the term "recombinant nucleic acid" herein is meant nucleic acid, originally formed in vitro, in general, by the manipulation of nucleic acid by polymerases and endonucleases, in a form not normally found in nature. Thus an isolated nucleic acid, in a linear form, or an expression vector formed in vitro by ligating DNA molecules that are not normally joined, are both considered recombinant for the purposes of this invention. It is understood that once a recombinant nucleic acid is made and reintroduced into a host cell or organism, it will replicate non-recombinantly, i.e. using the in vivo cellular machinery of the

host cell rather than in vitro manipulations; however, such nucleic acids, once produced recombinantly, although subsequently replicated non-recombinantly, are still considered recombinant for the purposes of the invention.

[29] Similarly, a "recombinant protein" is a protein made using recombinant techniques, i.e. through the expression of a recombinant nucleic acid as depicted above. A recombinant protein is distinguished from naturally occurring protein by at least one or more characteristics. For example, the protein may be isolated or purified away from some or all of the proteins and compounds with which it is normally associated in its wild type host, and thus may be substantially pure. For example, an isolated protein is unaccompanied by at least some of the material with which it is normally associated in its natural state, preferably constituting at least about 0.5%, more preferably at least about 5% by weight of the total protein in a given sample. A substantially pure protein comprises at least about 75% by weight of the total protein, with at least about 80% being preferred, and at least about 90% being particularly preferred. The definition includes the production of a colorectal cancer protein from one organism in a different organism or host cell. Alternatively, the protein may be made at a significantly higher concentration than is normally seen, through the use of an inducible promoter or high expression promoter, such that the protein is made at increased concentration levels. Alternatively, the protein may be in a form not normally found in nature, as in the addition of an epitope tag or amino acid substitutions, insertions and deletions, as discussed below.

[30] In a preferred embodiment, the colorectal cancer sequences are nucleic acids. As will be appreciated by those in the art and is more fully outlined below, colorectal cancer sequences are useful in a variety of applications, including diagnostic applications, which will detect naturally occurring nucleic acids, as well as screening applications; for example, biochips comprising nucleic acid probes to the colorectal cancer sequences can be generated. In the broadest sense, then, by "nucleic acid" or "oligonucleotide" or grammatical equivalents herein means at least two nucleotides covalently linked together. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, as outlined below, nucleic acid analogs are included that may have alternate backbones, comprising, for example, phosphoramidate (Beaucage et al., Tetrahedron 49(10):1925 (1993) and references therein; Letsinger, J. Org. Chem. 35:3800 (1970); Sprinzl et al., Eur. J. Biochem. 81:579 (1977); Letsinger et al., Nucl. Acids Res. 14:3487 (1986); Sawai et al, Chem. Lett. 805 (1984), Letsinger et al., J. Am. Chem. Soc. 110:4470 (1988); and Pauwels et al., Chemica Scripta 26:141 (1986)),

phosphorothioate (Mag et al., *Nucleic Acids Res.* 19:1437 (1991); and U.S. Patent No. 5,644,048), phosphorodithioate (Briu et al., *J. Am. Chem. Soc.* 111:2321 (1989), O-methylphosphoroamidite linkages (see Eckstein, *Oligonucleotides and Analogues: A Practical Approach*, Oxford University Press), and peptide nucleic acid backbones and linkages (see Egholm, *J. Am. Chem. Soc.* 114:1895 (1992); Meier et al., *Chem. Int. Ed. Engl.* 31:1008 (1992); Nielsen, *Nature*, 365:566 (1993); Carlsson et al., *Nature* 380:207 (1996), all of which are incorporated by reference). Other analog nucleic acids include those with positive backbones (Denpcy et al., *Proc. Natl. Acad. Sci. USA* 92:6097 (1995); non-ionic backbones (U.S. Patent Nos. 5,386,023, 5,637,684, 5,602,240, 5,216,141 and 4,469,863; Kiedrowski et al., *Angew. Chem. Intl. Ed. English* 30:423 (1991); Letsinger et al., *J. Am. Chem. Soc.* 110:4470 (1988); Letsinger et al., *Nucleoside & Nucleotide* 13:1597 (1994); Chapters 2 and 3, *ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research"*, Ed. Y.S. Sanghui and P. Dan Cook; Mesmaeker et al., *Bioorganic & Medicinal Chem. Lett.* 4:395 (1994); Jeffs et al., *J. Biomolecular NMR* 34:17 (1994); *Tetrahedron Lett.* 37:743 (1996)) and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, *ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research"*, Ed. Y.S. Sanghui and P. Dan Cook. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids (see Jenkins et al., *Chem. Soc. Rev.* (1995) pp169-176). Several nucleic acid analogs are described in Rawls, *C & E News* June 2, 1997 page 35. All of these references are hereby expressly incorporated by reference. These modifications of the ribose-phosphate backbone may be done for a variety of reasons, for example to increase the stability and half-life of such molecules in physiological environments or as probes on a biochip.

[31] As will be appreciated by those in the art, all of these nucleic acid analogs may find use in the present invention. In addition, mixtures of naturally occurring nucleic acids and analogs can be made; alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made.

[32] Particularly preferred are peptide nucleic acids (PNA) which includes peptide nucleic acid analogs. These backbones are substantially non-ionic under neutral conditions, in contrast to the highly charged phosphodiester backbone of naturally occurring nucleic acids. This results in two advantages. First, the PNA backbone exhibits improved hybridization kinetics. PNAs have larger changes in the melting temperature ( $T_m$ ) for mismatched versus perfectly matched basepairs. DNA and RNA typically exhibit a 2-4°C drop in  $T_m$  for an internal mismatch. With the non-ionic PNA backbone, the drop is closer to

7-9°C. Similarly, due to their non-ionic nature, hybridization of the bases attached to these backbones is relatively insensitive to salt concentration. In addition, PNAs are not degraded by cellular enzymes, and thus can be more stable.

[33] The nucleic acids may be single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. As will be appreciated by those in the art, the depiction of a single strand (“Watson”) also defines the sequence of the other strand (“Crick”); thus the sequences described herein also includes the complement of the sequence. The nucleic acid may be DNA, both genomic and cDNA, RNA or a hybrid, where the nucleic acid contains any combination of deoxyribo- and ribo-nucleotides, and any combination of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine hypoxanthine, isocytosine, isoguanine, etc. As used herein, the term “nucleoside” includes nucleotides and nucleoside and nucleotide analogs, and modified nucleosides such as amino modified nucleosides. In addition, “nucleoside” includes non-naturally occurring analog structures. Thus for example the individual units of a peptide nucleic acid, each containing a base, are referred to herein as a nucleoside.

[34] A colorectal cancer sequence can be initially identified by substantial nucleic acid and/or amino acid sequence homology to the colorectal cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either homology programs or hybridization conditions.

[35] The isolation of mRNA comprises isolating total cellular RNA by disrupting a cell and performing differential centrifugation. Once the total RNA is isolated, mRNA is isolated by making use of the adenine nucleotide residues known to those skilled in the art as a poly (A) tail found on virtually every eukaryotic mRNA molecule at the 3'end thereof. Oligonucleotides composed of only deoxythymidine [oligo(dT)] are linked to cellulose and the oligo(dT)-cellulose packed into small columns. When a preparation of total cellular RNA is passed through such a column, the mRNA molecules bind to the oligo(dT) by the poly (A) tails while the rest of the RNA flows through the column. The bound mRNAs are then eluted from the column and collected.

[36] The colorectal cancer sequences of the invention can be identified as follows. Samples of normal and tumor tissue are applied to biochips comprising nucleic acid probes. The samples are first microdissected, if applicable, and treated as described above for the preparation of mRNA. Suitable biochips are commercially available, for example

from Affymetrix. Gene expression profiles as described herein are generated, and the data analyzed.

[37] In a preferred embodiment, the genes showing changes in expression as between normal and disease states are compared to genes expressed in other normal tissues, including, but not limited to lung, heart, brain, liver, breast, kidney, muscle, prostate, small intestine, large intestine, spleen, bone, and placenta. In a preferred embodiment, those genes identified during the colorectal cancer screen that are expressed in any significant amount in other tissues are removed from the profile, although in some embodiments, this is not necessary. That is, when screening for drugs, it is preferable that the target be disease specific, to minimize possible side effects.

[38] In a preferred embodiment, colorectal cancer sequences are those that are up-regulated in colorectal cancer ; that is, the expression of these genes is higher in colorectal carcinoma as compared to normal colon tissue. "Up-regulation" as used herein means at least about a 1.1 fold change, preferably a 1.5 or two fold change, preferably at least about a three fold change, with at least about five-fold or higher being preferred. All accession numbers herein are for the GenBank sequence database and the sequences of the accession numbers are hereby expressly incorporated by reference. GenBank is known in the art, see, e.g., Benson, DA, et al., Nucleic Acids Research 26:1-7 (1998) and <http://www.ncbi.nlm.nih.gov/>. In addition, these genes were found to be expressed in a limited amount or not at all in heart, brain, lung, liver, breast, kidney, prostate, small intestine and spleen.

[39] In a preferred embodiment, colorectal cancer sequences are those that are down-regulated in colorectal cancer ; that is, the expression of these genes is lower in colorectal carcinoma as compared to normal colon tissue. "Down-regulation" as used herein means at least about a two-fold change, preferably at least about a three fold change, with at least about five-fold or higher being preferred.

[40] Colorectal cancer proteins of the present invention may be classified as secreted proteins, transmembrane proteins or intracellular proteins. In a preferred embodiment the colorectal cancer protein is an intracellular protein. Intracellular proteins may be found in the cytoplasm and/or in the nucleus. Intracellular proteins are involved in all aspects of cellular function and replication (including, for example, signaling pathways); aberrant expression of such proteins results in unregulated or disregulated cellular processes. For example, many intracellular proteins have enzymatic activity such as protein kinase activity, protein phosphatase activity, protease activity, nucleotide cyclase activity,

polymerase activity and the like. Intracellular proteins also serve as docking proteins that are involved in organizing complexes of proteins, or targeting proteins to various subcellular localizations, and are involved in maintaining the structural integrity of organelles.

[41] An increasingly appreciated concept in characterizing intracellular proteins is the presence in the proteins of one or more motifs for which defined functions have been attributed. In addition to the highly conserved sequences found in the enzymatic domain of proteins, highly conserved sequences have been identified in proteins that are involved in protein-protein interaction. For example, Src-homology-2 (SH2) domains bind tyrosine-phosphorylated targets in a sequence dependent manner. PTB domains, which are distinct from SH2 domains, also bind tyrosine phosphorylated targets. SH3 domains bind to proline-rich targets. In addition, PH domains, tetratricopeptide repeats and WD domains to name only a few, have been shown to mediate protein-protein interactions. Some of these may also be involved in binding to phospholipids or other second messengers. As will be appreciated by one of ordinary skill in the art, these motifs can be identified on the basis of primary sequence; thus, an analysis of the sequence of proteins may provide insight into both the enzymatic potential of the molecule and/or molecules with which the protein may associate.

[42] In a preferred embodiment, the colorectal cancer sequences are transmembrane proteins. Transmembrane proteins are molecules that span the phospholipid bilayer of a cell. They may have an intracellular domain, an extracellular domain, or both. The intracellular domains of such proteins may have a number of functions including those already described for intracellular proteins. For example, the intracellular domain may have enzymatic activity and/or may serve as a binding site for additional proteins. Frequently the intracellular domain of transmembrane proteins serves both roles. For example certain receptor tyrosine kinases have both protein kinase activity and SH2 domains. In addition, autophosphorylation of tyrosines on the receptor molecule itself, creates binding sites for additional SH2 domain containing proteins.

[43] Transmembrane proteins may contain from one to many transmembrane domains. For example, receptor tyrosine kinases, certain cytokine receptors, receptor guanylyl cyclases and receptor serine/threonine protein kinases contain a single transmembrane domain. However, various other proteins including channels and adenylyl cyclases contain numerous transmembrane domains. Many important cell surface receptors are classified as "seven transmembrane domain" proteins, as they contain 7 membrane spanning regions. Important transmembrane protein receptors include, but are not limited to

insulin receptor, insulin-like growth factor receptor, human growth hormone receptor, glucose transporters, transferrin receptor, epidermal growth factor receptor, low density lipoprotein receptor, epidermal growth factor receptor, leptin receptor, interleukin receptors, e.g. IL-1 receptor, IL-2 receptor, etc.

5                   [44]    Characteristics of transmembrane domains include approximately 20 consecutive hydrophobic amino acids that may be followed by charged amino acids. Therefore, upon analysis of the amino acid sequence of a particular protein, the localization and number of transmembrane domains within the protein may be predicted.

10                   [45]    The extracellular domains of transmembrane proteins are diverse; however, conserved motifs are found repeatedly among various extracellular domains. Conserved structure and/or functions have been ascribed to different extracellular motifs. For example, cytokine receptors are characterized by a cluster of cysteines and a WSXWS (W= tryptophan, S= serine, X=any amino acid) motif. Immunoglobulin-like domains are highly conserved. Mucin-like domains may be involved in cell adhesion and leucine-rich repeats  
15 participate in protein-protein interactions.

                  [46]    Many extracellular domains are involved in binding to other molecules. In one aspect, extracellular domains are receptors. Factors that bind the receptor domain include circulating ligands, which may be peptides, proteins, or small molecules such as adenosine and the like. For example, growth factors such as EGF, FGF and PDGF are  
20 circulating growth factors that bind to their cognate receptors to initiate a variety of cellular responses. Other factors include cytokines, mitogenic factors, neurotrophic factors and the like. Extracellular domains also bind to cell-associated molecules. In this respect, they mediate cell-cell interactions. Cell-associated ligands can be tethered to the cell for example via a glycosylphosphatidylinositol (GPI) anchor, or may themselves be transmembrane  
25 proteins. Extracellular domains also associate with the extracellular matrix and contribute to the maintenance of the cell structure.

                  [47]    Colorectal cancer proteins that are transmembrane are particularly preferred in the present invention as they are good targets for immunotherapeutics, as are described herein. In addition, as outlined below, transmembrane proteins can be also useful  
30 in imaging modalities.

                  [48]    It will also be appreciated by those in the art that a transmembrane protein can be made soluble by removing transmembrane sequences, for example through recombinant methods. Furthermore, transmembrane proteins that have been made soluble

can be made to be secreted through recombinant means by adding an appropriate signal sequence.

[49] In a preferred embodiment, the colorectal cancer proteins are secreted proteins; the secretion of which can be either constitutive or regulated. These proteins have a signal peptide or signal sequence that targets the molecule to the secretory pathway. Secreted proteins are involved in numerous physiological events; by virtue of their circulating nature, they serve to transmit signals to various other cell types. The secreted protein may function in an autocrine manner (acting on the cell that secreted the factor), a paracrine manner (acting on cells in close proximity to the cell that secreted the factor) or an endocrine manner (acting on cells at a distance). Thus secreted molecules find use in modulating or altering numerous aspects of physiology. colorectal cancer proteins that are secreted proteins are particularly preferred in the present invention as they serve as good targets for diagnostic markers, for example for blood tests.

[50] A colorectal cancer sequence is initially identified by substantial nucleic acid and/or amino acid sequence homology to the colorectal cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either homology programs or hybridization conditions.

[51] As used herein, the terms "colorectal cancer nucleic acid", "colorectal cancer protein" or "colorectal cancer polynucleotide" or "colorectal cancer-associated transcript" refers to nucleic acid and polypeptide polymorphic variants, alleles, mutants, and interspecies homologs that: (1) have a nucleotide sequence that has greater than about 60% nucleotide sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater nucleotide sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more nucleotides, to a nucleotide sequence of or associated with a unigene cluster of Tables 1 or Table 2; (2) bind to antibodies, e.g., polyclonal antibodies, raised against an immunogen comprising an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Table 1 or Table 2, and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to a nucleic acid sequence, or the complement thereof of Table 1 or Table 2 and conservatively modified variants thereof or (4) have an amino acid sequence that has greater than about 60% amino acid sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater amino sequence identity, preferably over a region of over a region of at least about

25, 50, 100, 200, 500, 1000, or more amino acid, to an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Table 1 or Table 2. A polynucleotide or polypeptide sequence is typically from a mammal including, but not limited to, primate, e.g., human; rodent, e.g., rat, mouse, hamster; cow, pig, horse, sheep, or  
5 other mammal. A “colorectal cancer polypeptide” and a “colorectal cancer polynucleotide,” include both naturally occurring or recombinant.

[52] Homology in this context means sequence similarity or identity, with identity being preferred. A preferred comparison for homology purposes is to compare the sequence containing sequencing errors to the correct sequence. This homology will be  
10 determined using standard techniques known in the art, including, but not limited to, the local homology algorithm of Smith & Waterman, Adv. Appl. Math. 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, J. Mol. Biol. 48:443 (1970), by the search for similarity method of Pearson & Lipman, PNAS USA 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA  
15 in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Drive, Madison, WI), the Best Fit sequence program described by Devereux et al., Nucl. Acid Res. 12:387-395 (1984), preferably using the default settings, or by inspection.

[53] In a preferred embodiment, the sequences which are used to determine sequence identity or similarity are selected from the sequences set forth in Table 1 or Table 2.  
20 In one embodiment the sequences utilized herein are those set forth in Table 1 or Table 2. In another embodiment, the sequences are naturally occurring allelic variants of the sequences set forth in Table 1 or Table 2. In another embodiment, the sequences are sequence variants as further described herein.

[54] The terms “identical” or percent “identity,” in the context of two or  
25 more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (i.e., about 60% identity, preferably 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region, when compared and aligned for maximum correspondence over a comparison window or designated region)  
30 as measured using a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (*see, e.g.*, NCBI web site <http://www.ncbi.nlm.nih.gov/BLAST/> or the like). Such sequences are then said to be “substantially identical.” This definition also refers to, or may be applied to, the complement of a test sequence. The definition also includes sequences that have deletions

and/or additions, as well as those that have substitutions, as well as naturally occurring, e.g., polymorphic or allelic variants, and man-made variants. As described below, the preferred algorithms can account for gaps and the like. Preferably, identity exists over a region that is at least about 25 amino acids or nucleotides in length, or more preferably over a region that is 50-100 amino acids or nucleotides in length.

[55] For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

[56] A "comparison window", as used herein, includes reference to a segment of one of the number of contiguous positions selected from the group consisting typically of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (*see, e.g., Current Protocols in Molecular Biology* (Ausubel *et al.*, eds. 1995 supplement)).

[57] Preferred examples of algorithms that are suitable for determining percent sequence identity and sequence similarity include the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.*, *Nuc. Acids Res.* 25:3389-3402 (1997) and Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990). BLAST and BLAST 2.0 are used, with the parameters described herein, to determine percent sequence identity for the nucleic acids and proteins of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>).

This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length  $W$  in the query sequence, which either match or satisfy some positive-valued threshold score  $T$  when aligned with a word of the same length in a database sequence.  $T$  is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*).

5 These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, e.g., for nucleotide sequences, the parameters  $M$  (reward score for a pair of matching residues; always  $> 0$ ) and  $N$  (penalty score for mismatching residues; always  $< 0$ ). For amino  
10 acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity  $X$  from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters  $W$ ,  $T$ , and  $X$  determine the  
15 sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength ( $W$ ) of 11, an expectation ( $E$ ) of 10,  $M=5$ ,  $N=-4$  and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation ( $E$ ) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alignments ( $B$ ) of  
20 50, expectation ( $E$ ) of 10,  $M=5$ ,  $N=-4$ , and a comparison of both strands.

[58] The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability ( $P(N)$ ), which provides an indication of the probability by which a  
25 match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001. Log values may be large negative numbers, e.g., 5, 10, 20, 30, 40, 40, 70, 90, 110, 150, 170,  
30 etc.

[59] In one embodiment, the nucleic acid homology is determined through hybridization studies. Thus, for example, nucleic acids which hybridize under high stringency to the nucleic acid sequences which encode the peptides identified in Table 1 or Table 2, or their complements, are considered a colorectal cancer sequence. High stringency

conditions are known in the art; see for example Maniatis et al., *Molecular Cloning: A Laboratory Manual*, 2d Edition, 1989, and *Short Protocols in Molecular Biology*, ed. Ausubel, et al., both of which are hereby incorporated by reference. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences  
5 hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are selected to be about 5-10°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a  
10 defined ionic strength pH. The  $T_m$  is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at  $T_m$ , 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M  
15 sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g. 10 to 50 nucleotides) and at least about 60°C for long probes (e.g. greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide.

[60] In another embodiment, less stringent hybridization conditions are  
20 used; for example, moderate or low stringency conditions may be used, as are known in the art; see Maniatis and Ausubel, *supra*, and Tijssen, *supra*. For selective or specific hybridization, a positive signal is at least two times background, preferably 10 times background hybridization. Exemplary stringent hybridization conditions can be as following: 50% formamide, 5x SSC, and 1% SDS, incubating at 42°C, or, 5x SSC, 1% SDS, incubating  
25 at 65°C, with wash in 0.2x SSC, and 0.1% SDS at 65°C.

[61] Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, for example, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the  
30 nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 1X SSC at 45°C. A positive hybridization is at least twice background. Those of ordinary skill will readily

recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous reference, e.g., and *Current Protocols in Molecular Biology*, ed. Ausubel, *et al.*

5                   [62] For PCR, a temperature of about 36°C is typical for low stringency amplification, although annealing temperatures may vary between about 32°C and 48°C depending on primer length. For high stringency PCR amplification, a temperature of about 62°C is typical, although high stringency annealing temperatures can range from about 50°C to about 65°C, depending on the primer length and specificity. Typical cycle conditions for  
10 both high and low stringency amplifications include a denaturation phase of 90°C - 95°C for 30 sec - 2 min., an annealing phase lasting 30 sec. - 2 min., and an extension phase of about 72°C for 1 - 2 min. Protocols and guidelines for low and high stringency amplification reactions are provided, e.g., in Innis *et al.*, *PCR Protocols, A Guide to Methods and Applications* (1990).

15                   [63] In addition, the colorectal cancer nucleic acid sequences of the invention are fragments of larger genes, i.e. they are nucleic acid segments. "Genes" in this context includes coding regions, non-coding regions, and mixtures of coding and non-coding regions. Accordingly, as will be appreciated by those in the art, using the sequences provided herein, additional sequences of the colorectal cancer genes can be obtained, using techniques  
20 well known in the art for cloning either longer sequences or the full length sequences; see Maniatis *et al.*, and Ausubel, *et al.*, *supra*, hereby expressly incorporated by reference.

                  [64] An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by  
25 the second nucleic acid. Thus, a polypeptide is typically substantially identical to a second polypeptide, e.g., where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions, as described above. Yet another indication that two nucleic acid sequences are substantially identical is that the  
30 same primers can be used to amplify the sequences.

                  [65] Once the colorectal cancer nucleic acid is identified, it can be cloned and, if necessary, its constituent parts recombined to form the entire colorectal cancer nucleic acid. Once isolated from its natural source, e.g., contained within a plasmid or other vector

or excised therefrom as a linear nucleic acid segment, the recombinant colorectal cancer nucleic acid can be further-used as a probe to identify and isolate other colorectal cancer nucleic acids, for example additional coding regions. It can also be used as a "precursor" nucleic acid to make modified or variant colorectal cancer nucleic acids and proteins.

5                   [66]    The colorectal cancer nucleic acids of the present invention are used in several ways. In a first embodiment, nucleic acid probes to the colorectal cancer nucleic acids are made and attached to biochips to be used in screening and diagnostic methods, as outlined below, or for administration, for example for gene therapy and/or antisense applications. Alternatively, the colorectal cancer nucleic acids that include coding regions of  
10 colorectal cancer proteins can be put into expression vectors for the expression of colorectal cancer proteins, again either for screening purposes or for administration to a patient.

                  [67]    In a preferred embodiment, nucleic acid probes to colorectal cancer nucleic acids (both the nucleic acid sequences encoding peptides outlined in the Table 1 or Table 2 and/or the complements thereof) are made. The nucleic acid probes attached to the  
15 biochip are designed to be substantially complementary to the colorectal cancer nucleic acids, i.e. the target sequence (either the target sequence of the sample or to other probe sequences, for example in sandwich assays), such that hybridization of the target sequence and the probes of the present invention occurs. As outlined below, this complementarity need not be perfect; there may be any number of base pair mismatches which will interfere with  
20 hybridization between the target sequence and the single stranded nucleic acids of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by "substantially complementary" herein is meant that the probes are sufficiently complementary to the target sequences to hybridize under  
25 normal reaction conditions, particularly high stringency conditions, as outlined herein.

                  [68]    A nucleic acid probe is generally single stranded but can be partially single and partially double stranded. The strandedness of the probe is dictated by the structure, composition, and properties of the target sequence. In general, the nucleic acid probes range from about 8 to about 100 bases long, with from about 10 to about 80 bases  
30 being preferred, and from about 30 to about 50 bases being particularly preferred. That is, generally whole genes are not used. In some embodiments, much longer nucleic acids can be used, up to hundreds of bases.

                  [69]    In a preferred embodiment, more than one probe per sequence is used, with either overlapping probes or probes to different sections of the target being used. That

is, two, three, four or more probes, with three being preferred, are used to build in a redundancy for a particular target. The probes can be overlapping (i.e. have some sequence in common), or separate.

[70] As will be appreciated by those in the art, nucleic acids can be attached or immobilized to a solid support in a wide variety of ways. By “immobilized” and grammatical equivalents herein is meant the association or binding between the nucleic acid probe and the solid support is sufficient to be stable under the conditions of binding, washing, analysis, and removal as outlined below. The binding can be covalent or non-covalent. By “non-covalent binding” and grammatical equivalents herein is meant one or more of either electrostatic, hydrophilic, and hydrophobic interactions. Included in non-covalent binding is the covalent attachment of a molecule, such as, streptavidin to the support and the non-covalent binding of the biotinylated probe to the streptavidin. By “covalent binding” and grammatical equivalents herein is meant that the two moieties, the solid support and the probe, are attached by at least one bond, including sigma bonds, pi bonds and coordination bonds. Covalent bonds can be formed directly between the probe and the solid support or can be formed by a cross linker or by inclusion of a specific reactive group on either the solid support or the probe or both molecules. Immobilization may also involve a combination of covalent and non-covalent interactions.

[71] In general, the probes are attached to the biochip in a wide variety of ways, as will be appreciated by those in the art. As described herein, the nucleic acids can either be synthesized first, with subsequent attachment to the biochip, or can be directly synthesized on the biochip.

[72] The biochip comprises a suitable solid substrate. By “substrate” or “solid support” or other grammatical equivalents herein is meant any material that can be modified to contain discrete individual sites appropriate for the attachment or association of the nucleic acid probes and is amenable to at least one detection method. As will be appreciated by those in the art, the number of possible substrates are very large, and include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, Teflon, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, etc. In general, the substrates allow optical detection and do not appreciably fluoresce. A preferred substrate is described in copending application entitled

Reusable Low Fluorescent Plastic Biochip, U.S. Application Serial No. 09/270,214, filed March 15, 1999, herein incorporated by reference in its entirety.

[73] Generally the substrate is planar, although as will be appreciated by those in the art, other configurations of substrates may be used as well. For example, the probes may be placed on the inside surface of a tube, for flow-through sample analysis to minimize sample volume. Similarly, the substrate may be flexible, such as a flexible foam, including closed cell foams made of particular plastics.

[74] In a preferred embodiment, the surface of the biochip and the probe may be derivatized with chemical functional groups for subsequent attachment of the two. Thus, for example, the biochip is derivatized with a chemical functional group including, but not limited to, amino groups, carboxy groups, oxo groups and thiol groups, with amino groups being particularly preferred. Using these functional groups, the probes can be attached using functional groups on the probes. For example, nucleic acids containing amino groups can be attached to surfaces comprising amino groups, for example using linkers as are known in the art; for example, homo-or hetero-bifunctional linkers as are well known (see 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200, incorporated herein by reference). In addition, in some cases, additional linkers, such as alkyl groups (including substituted and heteroalkyl groups) may be used.

[75] In this embodiment, the oligonucleotides are synthesized as is known in the art, and then attached to the surface of the solid support. As will be appreciated by those skilled in the art, either the 5' or 3' terminus may be attached to the solid support, or attachment may be via an internal nucleoside.

[76] In an additional embodiment, the immobilization to the solid support may be very strong, yet non-covalent. For example, biotinylated oligonucleotides can be made, which bind to surfaces covalently coated with streptavidin, resulting in attachment.

[77] Alternatively, the oligonucleotides may be synthesized on the surface, as is known in the art. For example, photoactivation techniques utilizing photopolymerization compounds and techniques are used. In a preferred embodiment, the nucleic acids can be synthesized in situ, using well known photolithographic techniques, such as those described in WO 95/25116; WO 95/35505; U.S. Patent Nos. 5,700,637 and 5,445,934; and references cited within, all of which are expressly incorporated by reference; these methods of attachment form the basis of the Affimetrix GeneChip™ technology.

[78] In a preferred embodiment, colorectal cancer nucleic acids encoding colorectal cancer proteins are used to make a variety of expression vectors to express colorectal cancer proteins which can then be used in screening assays, as described below. The expression vectors may be either self-replicating extrachromosomal vectors or vectors which integrate into a host genome. Generally, these expression vectors include transcriptional and translational regulatory nucleic acid operably linked to the nucleic acid encoding the colorectal cancer protein. The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

[79] Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice. The transcriptional and translational regulatory nucleic acid will generally be appropriate to the host cell used to express the colorectal cancer protein; for example, transcriptional and translational regulatory nucleic acid sequences from *Bacillus* are preferably used to express the colorectal cancer protein in *Bacillus*. Numerous types of appropriate expression vectors, and suitable regulatory sequences are known in the art for a variety of host cells.

[80] In general, the transcriptional and translational regulatory sequences may include, but are not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. In a preferred embodiment, the regulatory sequences include a promoter and transcriptional start and stop sequences.

[81] Promoter sequences encode either constitutive or inducible promoters. The promoters may be either naturally occurring promoters or hybrid promoters. Hybrid

promoters, which combine elements of more than one promoter, are also known in the art, and are useful in the present invention.

[82] In addition, the expression vector may comprise additional elements. For example, the expression vector may have two replication systems, thus allowing it to be maintained in two organisms, for example in mammalian or insect cells for expression and in a procaryotic host for cloning and amplification. Furthermore, for integrating expression vectors, the expression vector contains at least one sequence homologous to the host cell genome, and preferably two homologous sequences which flank the expression construct. The integrating vector may be directed to a specific locus in the host cell by selecting the appropriate homologous sequence for inclusion in the vector. Constructs for integrating vectors are well known in the art.

[83] In addition, in a preferred embodiment, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes are well known in the art and will vary with the host cell used.

[84] The colorectal cancer proteins of the present invention are produced by culturing a host cell transformed with an expression vector containing nucleic acid encoding a colorectal cancer protein, under the appropriate conditions to induce or cause expression of the colorectal cancer protein. The conditions appropriate for colorectal cancer protein expression will vary with the choice of the expression vector and the host cell, and will be easily ascertained by one skilled in the art through routine experimentation. For example, the use of constitutive promoters in the expression vector will require optimizing the growth and proliferation of the host cell, while the use of an inducible promoter requires the appropriate growth conditions for induction. In addition, in some embodiments, the timing of the harvest is important. For example, the baculoviral systems used in insect cell expression are lytic viruses, and thus harvest time selection can be crucial for product yield.

[85] Appropriate host cells include yeast, bacteria, archaebacteria, fungi, and insect and animal cells, including mammalian cells. Of particular interest are *Drosophila melanogaster* cells, *Saccharomyces cerevisiae* and other yeasts, *E. coli*, *Bacillus subtilis*, Sf9 cells, C129 cells, 293 cells, *Neurospora*, BHK, CHO, COS, HeLa cells, THP1 cell line (a macrophage cell line) and human cells and cell lines.

[86] In a preferred embodiment, the colorectal cancer proteins are expressed in mammalian cells. Mammalian expression systems are also known in the art, and include retroviral systems. A preferred expression vector system is a retroviral vector system such as is generally described in PCT/US97/01019 and PCT/US97/01048, both of which are

hereby expressly incorporated by reference. Of particular use as mammalian promoters are the promoters from mammalian viral genes, since the viral genes are often highly expressed and have a broad host range. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter, herpes simplex virus promoter, and the CMV promoter. Typically, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. Examples of transcription terminator and polyadenylation signals include those derived from SV40.

[87] The methods of introducing exogenous nucleic acid into mammalian hosts, as well as other hosts, is well known in the art, and will vary with the host cell used. Techniques include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, viral infection, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

[88] In a preferred embodiment, colorectal cancer proteins are expressed in bacterial systems. Bacterial expression systems are well known in the art. Promoters from bacteriophage may also be used and are known in the art. In addition, synthetic promoters and hybrid promoters are also useful; for example, the tac promoter is a hybrid of the trp and lac promoter sequences. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. In addition to a functioning promoter sequence, an efficient ribosome binding site is desirable. The expression vector may also include a signal peptide sequence that provides for secretion of the colorectal cancer protein in bacteria. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). The bacterial expression vector may also include a selectable marker gene to allow for the selection of bacterial strains that have been transformed. Suitable selection genes include genes which render the bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin and tetracycline. Selectable markers also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways. These components are assembled into expression vectors. Expression vectors for bacteria are well known in the art, and include vectors for *Bacillus subtilis*, *E. coli*, *Streptococcus cremoris*, and *Streptococcus lividans*, among others. The bacterial expression vectors are transformed

into bacterial host cells using techniques well known in the art, such as calcium chloride treatment, electroporation, and others.

[89] In one embodiment, colorectal cancer proteins are produced in insect cells. Expression vectors for the transformation of insect cells, and in particular, baculovirus-based expression vectors, are well known in the art.

[90] In a preferred embodiment, colorectal cancer protein is produced in yeast cells. Yeast expression systems are well known in the art, and include expression vectors for *Saccharomyces cerevisiae*, *Candida albicans* and *C. maltosa*, *Hansenula polymorpha*, *Kluyveromyces fragilis* and *K. lactis*, *Pichia guilliermondii* and *P. pastoris*, *Schizosaccharomyces pombe*, and *Yarrowia lipolytica*.

[91] The colorectal cancer protein may also be made as a fusion protein, using techniques well known in the art. Thus, for example, for the creation of monoclonal antibodies, if the desired epitope is small, the colorectal cancer protein may be fused to a carrier protein to form an immunogen. Alternatively, the colorectal cancer protein may be made as a fusion protein to increase expression, or for other reasons. For example, when the colorectal cancer protein is a colorectal cancer peptide, the nucleic acid encoding the peptide may be linked to other nucleic acid for expression purposes.

[92] In one embodiment, the colorectal cancer nucleic acids, proteins and antibodies of the invention are labeled. By "labeled" herein is meant that a compound has at least one element, isotope or chemical compound attached to enable the detection of the compound. In general, labels fall into three classes: a) isotopic labels, which may be radioactive or heavy isotopes; b) immune labels, which may be antibodies or antigens; and c) colored or fluorescent dyes. The labels may be incorporated into the colorectal cancer nucleic acids, proteins and antibodies at any position. For example, the label should be capable of producing, either directly or indirectly, a detectable signal. The detectable moiety may be a radioisotope, such as  $^3\text{H}$ ,  $^{14}\text{C}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ , or  $^{125}\text{I}$ , a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the label may be employed, including those methods described by Hunter et al., *Nature*, 144:945 (1962); David et al., *Biochemistry*, 13:1014 (1974); Pain et al., *J. Immunol. Meth.*, 40:219 (1981); and Nygren, *J. Histochem. and Cytochem.*, 30:407 (1982).

[93] Accordingly, the present invention also provides colorectal cancer protein sequences. A colorectal cancer protein of the present invention may be identified in

several ways. "Protein" in this sense includes proteins, polypeptides, and peptides terms which are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers, those containing modified residues, and non-naturally occurring amino acid polymer.

[94] As will be appreciated by those in the art, the nucleic acid sequences of the invention can be used to generate protein sequences. There are a variety of ways to do this, including cloning the entire gene and verifying its frame and amino acid sequence, or by comparing it to known sequences to search for homology to provide a frame, assuming the colorectal cancer protein has homology to some protein in the database being used. Generally, the nucleic acid sequences are input into a program that will search all three frames for homology. This is done in a preferred embodiment using the following NCBI Advanced BLAST parameters. The program is blastx or blastn. The database is nr. The input data is as "Sequence in FASTA format". The organism list is "none". The "expect" is 10; the filter is default. The "descriptions" is 500, the "alignments" is 500, and the "alignment view" is pairwise. The "Query Genetic Codes" is standard (1). The matrix is BLOSUM62; gap existence cost is 11, per residue gap cost is 1; and the lambda ratio is .85 default. This results in the generation of a putative protein sequence.

[95] Also included within one embodiment of colorectal cancer proteins are amino acid variants of the naturally occurring sequences, as determined herein. Preferably, the variants are preferably greater than about 75% homologous to the wild-type sequence, more preferably greater than about 80%, even more preferably greater than about 85% and most preferably greater than 90%. In some embodiments the homology will be as high as about 93 to 95 or 98%. As for nucleic acids, homology in this context means sequence similarity or identity, with identity being preferred. This homology will be determined using standard techniques known in the art as are outlined above for the nucleic acid homologies.

[96] Colorectal cancer proteins of the present invention may be shorter or longer than the wild type amino acid sequences. Thus, in a preferred embodiment, included within the definition of colorectal cancer proteins are portions or fragments of the wild type sequences. herein. In addition, as outlined above, the colorectal cancer nucleic acids of the invention may be used to obtain additional coding regions, and thus additional protein sequence, using techniques known in the art.

[97] In a preferred embodiment, the colorectal cancer proteins are derivative or variant colorectal cancer proteins as compared to the wild-type sequence. That is, as outlined more fully below, the derivative colorectal cancer peptide will contain at least one amino acid substitution, deletion or insertion, with amino acid substitutions being particularly preferred. The amino acid substitution, insertion or deletion may occur at any residue within the colorectal cancer peptide.

[98] Also included in an embodiment of colorectal cancer proteins of the present invention are amino acid sequence variants. These variants fall into one or more of three classes: substitutional, insertional or deletional variants. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the colorectal cancer protein, using cassette or PCR mutagenesis or other techniques well known in the art, to produce DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture as outlined above. However, variant colorectal cancer protein fragments having up to about 100-150 residues may be prepared by in vitro synthesis using established techniques. Amino acid sequence variants are characterized by the predetermined nature of the variation, a feature that sets them apart from naturally occurring allelic or interspecies variation of the colorectal cancer protein amino acid sequence. The variants typically exhibit the same qualitative biological activity as the naturally occurring analogue, although variants can also be selected which have modified characteristics as will be more fully outlined below.

[99] While the site or region for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region and the expressed colorectal cancer variants screened for the optimal combination of desired activity. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, for example, M13 primer mutagenesis and PCR mutagenesis. Screening of the mutants is done using assays of colorectal cancer protein activities.

[100] Amino acid substitutions are typically of single residues; insertions usually will be on the order of from about 1 to 20 amino acids, although considerably larger insertions may be tolerated. Deletions range from about 1 to about 20 residues, although in some cases deletions may be much larger.

[101] Substitutions, deletions, insertions or any combination thereof may be used to arrive at a final derivative. Generally these changes are done on a few amino acids to minimize the alteration of the molecule. However, larger changes may be tolerated in certain

circumstances. When small alterations in the characteristics of the colorectal cancer protein are desired, substitutions are generally made in accordance with the following chart:

**Chart I**

|    | Original Residue | Exemplary Substitutions |
|----|------------------|-------------------------|
| 5  | Ala              | Ser                     |
|    | Arg              | Lys                     |
|    | Asn              | Gln, His                |
|    | Asp              | Glu                     |
| 10 | Cys              | Ser                     |
|    | Gln              | Asn                     |
|    | Glu              | Asp                     |
|    | Gly              | Pro                     |
|    | His              | Asn, Gln                |
| 15 | Ile              | Leu, Val                |
|    | Leu              | Ile, Val                |
|    | Lys              | Arg, Gln, Glu           |
|    | Met              | Leu, Ile                |
|    | Phe              | Met, Leu, Tyr           |
| 20 | Ser              | Thr                     |
|    | Thr              | Ser                     |
|    | Trp              | Tyr                     |
|    | Tyr              | Trp, Phe                |
|    | Val              | Ile, Leu                |
| 25 |                  |                         |

[102] Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those shown in Chart I. For example, substitutions may be made which more significantly affect: the structure of the polypeptide backbone in the area of the alteration, for example the alpha-helical or beta-sheet structure; the charge or hydrophobicity of the molecule at the target site; or the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in the polypeptide's properties are those in which (a) a hydrophilic residue, e.g. seryl or threonyl is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue

having an electropositive side chain, e.g. lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g. glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g. phenylalanine, is substituted for (or by) one not having a side chain, e.g. glycine.

[103] The variants typically exhibit the same qualitative biological activity and will elicit the same immune response as the naturally-occurring analogue, although variants also are selected to modify the characteristics of the colorectal cancer proteins as needed. Alternatively, the variant may be designed such that the biological activity of the colorectal cancer protein is altered. For example, glycosylation sites may be altered or removed.

[104] Covalent modifications of colorectal cancer polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a colorectal cancer polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N-or C-terminal residues of a colorectal cancer polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking colorectal cancer to a water-insoluble support matrix or surface for use in the method for purifying anti-colorectal cancer antibodies or screening assays, as is more fully described below. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxy-succinimide esters, for example, esters with 4-azido-salicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis-(succinimidyl-propionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)-dithio]propionate.

[105] Other modifications include deamidation of glutamyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl, threonyl or tyrosyl residues, methylation of the  $\alpha$ -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, *Proteins: Structure and Molecular Properties*, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

[106] Another type of covalent modification of the colorectal cancer polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence colorectal cancer polypeptide, and/or adding one or more glycosylation sites that are not present in the native sequence colorectal cancer polypeptide.

[107] Addition of glycosylation sites to colorectal cancer polypeptides may be accomplished by altering the amino acid sequence thereof. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence colorectal cancer polypeptide (for O-linked glycosylation sites). The colorectal cancer amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the colorectal cancer polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

[108] Another means of increasing the number of carbohydrate moieties on the colorectal cancer polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, colorectal cancer Crit. Rev. Biochem., pp. 259-306 (1981).

[109] Removal of carbohydrate moieties present on the colorectal cancer polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo-and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

[110] Another type of covalent modification of colorectal cancer comprises linking the colorectal cancer polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

[111] colorectal cancer polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising a colorectal cancer polypeptide fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric molecule comprises a fusion of a colorectal cancer polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino-or carboxyl-terminus of the colorectal cancer polypeptide. The presence of such epitope-tagged forms of a colorectal cancer polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the colorectal cancer polypeptide to be readily purified by affinity purification

using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a colorectal cancer polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

[112] Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

[113] Also included with the definition of colorectal cancer protein in one embodiment are other colorectal cancer proteins of the colorectal cancer family, and colorectal cancer proteins from other organisms, which are cloned and expressed as outlined below. Thus, probe or degenerate polymerase chain reaction (PCR) primer sequences may be used to find other related colorectal cancer proteins from humans or other organisms. As will be appreciated by those in the art, particularly useful probe and/or PCR primer sequences include the unique areas of the colorectal cancer nucleic acid sequence. As is generally known in the art, preferred PCR primers are from about 15 to about 35 nucleotides in length, with from about 20 to about 30 being preferred, and may contain inosine as needed. The conditions for the PCR reaction are well known in the art.

[114] In addition, as is outlined herein, colorectal cancer proteins can be made that are longer than those depicted in the Table 1 or Table 2 for example, by the elucidation of additional sequences, the addition of epitope or purification tags, the addition of other fusion sequences, etc.

[115] Colorectal cancer proteins may also be identified as being encoded by colorectal cancer nucleic acids. Thus, colorectal cancer proteins are encoded by nucleic acids that will hybridize to the sequences of the sequence listings, or their complements, as outlined herein.

[116] In a preferred embodiment, when the colorectal cancer protein is to be used to generate antibodies, for example for immunotherapy, the colorectal cancer protein should share at least one epitope or determinant with the full length protein. By "epitope" or "determinant" herein is meant a portion of a protein which will generate and/or bind an antibody or T-cell receptor in the context of MHC. Thus, in most instances, antibodies made to a smaller colorectal cancer protein will be able to bind to the full length protein. In a preferred embodiment, the epitope is unique; that is, antibodies generated to a unique epitope show little or no cross-reactivity. In a preferred embodiment, the epitope is selected from a peptide encoded by a nucleic acid of Table 1. In another preferred embodiment, the epitope is selected from the CBF9 peptide sequence shown in Table 2.

[117] In one embodiment, the term "antibody" includes antibody fragments, as are known in the art, including Fab, Fab2, single chain antibodies (Fv for example), chimeric antibodies, etc., either produced by the modification of whole antibodies or those synthesized de novo using recombinant DNA technologies.

[118] Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the CBF9 peptide of Table 2, or a peptide encoded by a nucleic acid of Table 1 or fragment thereof or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

[119] The antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, *Nature*, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized in vitro. The immunizing agent will typically include the CBF9 polypeptide or a peptide encoded by a

nucleic acid of Table 1 or a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

[120] In one embodiment, the antibodies are bispecific antibodies.

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for a colorectal cancer protein or a fragment thereof, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit, preferably one that is tumor specific.

[121] In a preferred embodiment, the antibodies to colorectal cancer are capable of reducing or eliminating the biological function of colorectal cancer, as is described below. That is, the addition of anti-colorectal cancer antibodies (either polyclonal or preferably monoclonal) to colorectal cancer (or cells containing colorectal cancer) may reduce or eliminate the colorectal cancer activity. Generally, at least a 25% decrease in activity is preferred, with at least about 50% being particularly preferred and about a 95-100% decrease being especially preferred.

[122] In a preferred embodiment the antibodies to the colorectal cancer proteins are humanized antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric molecules of immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')<sub>2</sub> or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues form a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired

specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise  
5 substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et  
10 al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992)].

[123] Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred  
15 to as import residues, which are typically taken from an import variable domain.

Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-327 (1988); Verhoeven et al., *Science*, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR  
20 sequences for the corresponding sequences of a human antibody. Accordingly, such humanized antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

[124] Human antibodies can also be produced using various techniques  
25 known in the art, including phage display libraries [Hoogenboom and Winter, *J. Mol. Biol.*, 227:381 (1991); Marks et al., *J. Mol. Biol.*, 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, p. 77 (1985) and  
30 Boerner et al., *J. Immunol.*, 147(1):86-95 (1991)]. Similarly, human antibodies can be made by introducing of human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire.

This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks et al., *Bio/Technology* 10, 779-783 (1992); Lonberg et al., *Nature* 368 856-859 (1994); Morrison, *Nature* 368, 812-13 (1994); Fishwild et al., *Nature Biotechnology* 14, 845-51 (1996); Neuberger, *Nature Biotechnology* 14, 826 (1996); Lonberg and Huszar, *Intern. Rev. Immunol.* 13 65-93 (1995).

[125] By immunotherapy is meant treatment of colorectal cancer with an antibody raised against colorectal cancer proteins. As used herein, immunotherapy can be passive or active. Passive immunotherapy as defined herein is the passive transfer of antibody to a recipient (patient). Active immunization is the induction of antibody and/or T-cell responses in a recipient (patient). Induction of an immune response is the result of providing the recipient with an antigen to which antibodies are raised. As appreciated by one of ordinary skill in the art, the antigen may be provided by injecting a polypeptide against which antibodies are desired to be raised into a recipient, or contacting the recipient with a nucleic acid capable of expressing the antigen and under conditions for expression of the antigen.

[126] In a preferred embodiment the colorectal cancer proteins against which antibodies are raised are secreted proteins as described above. Without being bound by theory, antibodies used for treatment, bind and prevent the secreted protein from binding to its receptor, thereby inactivating the secreted colorectal cancer protein.

[127] In another preferred embodiment, the colorectal cancer protein to which antibodies are raised is a transmembrane protein. Without being bound by theory, antibodies used for treatment, bind the extracellular domain of the colorectal cancer protein and prevent it from binding to other proteins, such as circulating ligands or cell-associated molecules. The antibody may cause down-regulation of the transmembrane colorectal cancer protein. As will be appreciated by one of ordinary skill in the art, the antibody may be a competitive, non-competitive or uncompetitive inhibitor of protein binding to the extracellular domain of the colorectal cancer protein. The antibody is also an antagonist of the colorectal cancer protein. Further, the antibody prevents activation of the transmembrane colorectal cancer protein. In one aspect, when the antibody prevents the binding of other molecules to the colorectal cancer protein, the antibody prevents growth of the cell. The antibody also sensitizes the cell to cytotoxic agents, including, but not limited to TNF- $\alpha$ , TNF- $\beta$ , IL-1, INF- $\gamma$  and IL-2, or chemotherapeutic agents including 5FU, vinblastine,

actinomycin D, cisplatin, methotrexate, and the like. In some instances the antibody belongs to a sub-type that activates serum complement when complexed with the transmembrane protein thereby mediating cytotoxicity. Thus, colorectal cancer is treated by administering to a patient antibodies directed against the transmembrane colorectal cancer protein.

5           [128] In another preferred embodiment, the antibody is conjugated to a therapeutic moiety. In one aspect the therapeutic moiety is a small molecule that modulates the activity of the colorectal cancer protein. In another aspect the therapeutic moiety modulates the activity of molecules associated with or in close proximity to the colorectal cancer protein. The therapeutic moiety may inhibit enzymatic activity such as protease or  
10 protein kinase activity associated with colorectal cancer .

          [129] In a preferred embodiment, the therapeutic moiety may also be a cytotoxic agent. In this method, targeting the cytotoxic agent to tumor tissue or cells, results in a reduction in the number of afflicted cells, thereby reducing symptoms associated with colorectal cancer . Cytotoxic agents are numerous and varied and include, but are not limited to, cytotoxic drugs or toxins or active fragments of such toxins. Suitable toxins and their  
15 corresponding fragments include diptheria A chain, exotoxin A chain, ricin A chain, abrin A chain, curcin, crotin, phenomycin, enomycin and the like. Cytotoxic agents also include radiochemicals made by conjugating radioisotopes to antibodies raised against colorectal cancer proteins, or binding of a radionuclide to a chelating agent that has been covalently  
20 attached to the antibody. Targeting the therapeutic moiety to transmembrane colorectal cancer proteins not only serves to increase the local concentration of therapeutic moiety in the colorectal cancer afflicted area, but also serves to reduce deleterious side effects that may be associated with the therapeutic moiety.

          [130] In another preferred embodiment, the colorectal cancer protein against  
25 which the antibodies are raised is an intracellular protein. In this case, the antibody may be conjugated to a protein which facilitates entry into the cell. In one case, the antibody enters the cell by endocytosis. In another embodiment, a nucleic acid encoding the antibody is administered to the individual or cell. Moreover, wherein the colorectal cancer protein can be targeted within a cell, i.e., the nucleus, an antibody thereto contains a signal for that target  
30 localization, i.e., a nuclear localization signal.

          [131] The colorectal cancer antibodies of the invention specifically bind to colorectal cancer proteins. By "specifically bind" herein is meant that the antibodies bind to the protein with a binding constant in the range of at least  $10^{-4}$ -  $10^{-6}$   $M^{-1}$ , with a preferred range being  $10^{-7}$  -  $10^{-9}$   $M^{-1}$ .

[132] In a preferred embodiment, the colorectal cancer protein is purified or isolated after expression. Colorectal cancer proteins may be isolated or purified in a variety of ways known to those skilled in the art depending on what other components are present in the sample. Standard purification methods include electrophoretic, molecular,  
5 immunological and chromatographic techniques, including ion exchange, hydrophobic, affinity, and reverse-phase HPLC chromatography, and chromatofocusing. For example, the colorectal cancer protein may be purified using a standard anti-colorectal cancer antibody column. Ultrafiltration and diafiltration techniques, in conjunction with protein concentration, are also useful. For general guidance in suitable purification techniques, see  
10 Scopes, R., Protein Purification, Springer-Verlag, NY (1982). The degree of purification necessary will vary depending on the use of the colorectal cancer protein. In some instances no purification will be necessary.

[133] Once expressed and purified if necessary, the colorectal cancer proteins and nucleic acids are useful in a number of applications.

[134] In one aspect, the expression levels of genes are determined for  
15 different cellular states in the colorectal cancer phenotype; that is, the expression levels of genes in normal colon tissue and in colorectal cancer tissue (and in some cases, for varying severities of colorectal cancer that relate to prognosis, as outlined below) are evaluated to provide expression profiles. An expression profile of a particular cell state or point of  
20 development is essentially a "fingerprint" of the state; while two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is unique to the state of the cell. By comparing expression profiles of cells in different states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is  
25 obtained. Then, diagnosis may be done or confirmed: does tissue from a particular patient have the gene expression profile of normal or colorectal cancer tissue.

[135] "Differential expression," or grammatical equivalents as used herein, refers to both qualitative as well as quantitative differences in the genes' temporal and/or cellular expression patterns within and among the cells. Thus, a differentially expressed gene  
30 can qualitatively have its expression altered, including an activation or inactivation, in, for example, normal versus colorectal cancer tissue. That is, genes may be turned on or turned off in a particular state, relative to another state. As is apparent to the skilled artisan, any comparison of two or more states can be made. Such a qualitatively regulated gene will exhibit an expression pattern within a state or cell type which is detectable by standard

techniques in one such state or cell type, but is not detectable in both. Alternatively, the determination is quantitative in that expression is increased or decreased; that is, the expression of the gene is either upregulated, resulting in an increased amount of transcript, or downregulated, resulting in a decreased amount of transcript. The degree to which expression differs need only be large enough to quantify via standard characterization techniques as outlined below, such as by use of Affymetrix GeneChip™ expression arrays, Lockhart, Nature Biotechnology, 14:1675-1680 (1996), hereby expressly incorporated by reference. Other techniques include, but are not limited to, quantitative reverse transcriptase PCR, Northern analysis and RNase protection. As outlined above, preferably the change in expression (i.e. upregulation or downregulation) is at least about 50%, more preferably at least about 100%, more preferably at least about 150%, more preferably, at least about 200%, with from 300 to at least 1000% being especially preferred.

[136] As will be appreciated by those in the art, this may be done by evaluation at either the gene transcript, or the protein level; that is, the amount of gene expression may be monitored using nucleic acid probes to the DNA or RNA equivalent of the gene transcript, and the quantification of gene expression levels, or, alternatively, the final gene product itself (protein) can be monitored, for example through the use of antibodies to the colorectal cancer protein and standard immunoassays (ELISAs, etc.) or other techniques, including mass spectroscopy assays, 2D gel electrophoresis assays, etc. Thus, the proteins corresponding to colorectal cancer genes, i.e. those identified as being important in a colorectal cancer phenotype, can be evaluated in a colorectal cancer diagnostic test.

[137] In a preferred embodiment, gene expression monitoring is done and a number of genes, i.e. an expression profile, is monitored simultaneously, although multiple protein expression monitoring can be done as well. Similarly, these assays may be done on an individual basis as well.

[138] In this embodiment, the colorectal cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of colorectal cancer sequences in a particular cell. The assays are further described below in the example.

[139] In a preferred embodiment nucleic acids encoding the colorectal cancer protein are detected. Although DNA or RNA encoding the colorectal cancer protein may be detected, of particular interest are methods wherein the mRNA encoding a colorectal cancer protein is detected. The presence of mRNA in a sample is an indication that the colorectal cancer gene has been transcribed to form the mRNA, and suggests that the protein

is expressed. Probes to detect the mRNA can be any nucleotide/deoxynucleotide probe that is complementary to and base pairs with the mRNA and includes but is not limited to oligonucleotides, cDNA or RNA. Probes also should contain a detectable label, as defined herein. In one method the mRNA is detected after immobilizing the nucleic acid to be  
5 examined on a solid support such as nylon membranes and hybridizing the probe with the sample. Following washing to remove the non-specifically bound probe, the label is detected. In another method detection of the mRNA is performed in situ. In this method permeabilized cells or tissue samples are contacted with a detectably labeled nucleic acid probe for sufficient time to allow the probe to hybridize with the target mRNA. Following  
10 washing to remove the non-specifically bound probe, the label is detected. For example a digoxigenin labeled riboprobe (RNA probe) that is complementary to the mRNA encoding a colorectal cancer protein is detected by binding the digoxigenin with an anti-digoxigenin secondary antibody and developed with nitro blue tetrazolium and 5-bromo-4-chloro-3-indoyl phosphate.

15 [140] In a preferred embodiment, any of the three classes of proteins as described herein (secreted, transmembrane or intracellular proteins) are used in diagnostic assays. The colorectal cancer proteins, antibodies, nucleic acids, modified proteins and cells containing colorectal cancer sequences are used in diagnostic assays. This can be done on an individual gene or corresponding polypeptide level. In a preferred embodiment, the  
20 expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes and/or corresponding polypeptides.

[141] As described and defined herein, colorectal cancer proteins, including intracellular, transmembrane or secreted proteins, find use as markers of colorectal cancer .

25 Detection of these proteins in putative colorectal cancer tissue or patients allows for a determination or diagnosis of colorectal cancer . Numerous methods known to those of ordinary skill in the art find use in detecting colorectal cancer . In one embodiment, antibodies are used to detect colorectal cancer proteins. A preferred method separates proteins from a sample or patient by electrophoresis on a gel (typically a denaturing and  
30 reducing protein gel, but may be any other type of gel including isoelectric focusing gels and the like). Following separation of proteins, the colorectal cancer protein is detected by immunoblotting with antibodies raised against the colorectal cancer protein. Methods of immunoblotting are well known to those of ordinary skill in the art.

[142] In another preferred method, antibodies to the colorectal cancer protein find use in in situ imaging techniques. In this method cells are contacted with from one to many antibodies to the colorectal cancer protein(s). Following washing to remove non-specific antibody binding, the presence of the antibody or antibodies is detected. In one  
5 embodiment the antibody is detected by incubating with a secondary antibody that contains a detectable label. In another method the primary antibody to the colorectal cancer protein(s) contains a detectable label. In another preferred embodiment each one of multiple primary antibodies contains a distinct and detectable label. This method finds particular use in simultaneous screening for a plurality of colorectal cancer proteins. As will be appreciated  
10 by one of ordinary skill in the art, numerous other histological imaging techniques are useful in the invention.

[143] In a preferred embodiment the label is detected in a fluorometer which has the ability to detect and distinguish emissions of different wavelengths. In addition, a fluorescence activated cell sorter (FACS) can be used in the method.

[144] In another preferred embodiment, antibodies find use in diagnosing colorectal cancer from blood samples. As previously described, certain colorectal cancer proteins are secreted/circulating molecules. Blood samples, therefore, are useful as samples to be probed or tested for the presence of secreted colorectal cancer proteins. Antibodies can be used to detect the colorectal cancer by any of the previously described immunoassay  
20 techniques including ELISA, immunoblotting (Western blotting), immunoprecipitation, BIACORE technology and the like, as will be appreciated by one of ordinary skill in the art.

[145] In a preferred embodiment, in situ hybridization of labeled colorectal cancer nucleic acid probes to tissue arrays is done. For example, arrays of tissue samples, including colorectal cancer tissue and/or normal tissue, are made. In situ hybridization as is  
25 known in the art can then be done.

[146] It is understood that when comparing the fingerprints between an individual and a standard, the skilled artisan can make a diagnosis as well as a prognosis. It is further understood that the genes which indicate the diagnosis may differ from those which indicate the prognosis.

[147] In a preferred embodiment, the colorectal cancer proteins, antibodies, nucleic acids, modified proteins and cells containing colorectal cancer sequences are used in prognosis assays. As above, gene expression profiles can be generated that correlate to colorectal cancer severity, in terms of long term prognosis. Again, this may be done on either a protein or gene level, with the use of genes being preferred. As above, the colorectal  
30

cancer probes are attached to biochips for the detection and quantification of colorectal cancer sequences in a tissue or patient. The assays proceed as outlined for diagnosis.

[148] In a preferred embodiment, any of the three classes of proteins as described herein are used in drug screening assays. The colorectal cancer proteins, antibodies, nucleic acids, modified proteins and cells containing colorectal cancer sequences are used in drug screening assays or by evaluating the effect of drug candidates on a “gene expression profile” or expression profile of polypeptides. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent, Zlokarnik, et al., Science 279, 84-8 (1998), Heid, 1996 #69.

[149] In a preferred embodiment, the colorectal cancer proteins, antibodies, nucleic acids, modified proteins and cells containing the native or modified colorectal cancer proteins are used in screening assays. That is, the present invention provides novel methods for screening for compositions which modulate the colorectal cancer phenotype. As above, this can be done on an individual gene level or by evaluating the effect of drug candidates on a “gene expression profile”. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent, see Zlokarnik, supra.

Having identified the differentially expressed genes herein, a variety of assays may be executed. In a preferred embodiment, assays may be run on an individual gene or protein level. That is, having identified a particular gene as up regulated in colorectal cancer, candidate bioactive agents may be screened to modulate this gene’s response; preferably to down regulate the gene, although in some circumstances to up regulate the gene.

[150] The phrase “functional effects” in the context of assays for testing compounds that modulate activity of a colorectal cancer protein or colorectal cancer nucleic acid includes the determination of a parameter that is indirectly or directly under the influence of a colorectal cancer protein or nucleic acid, *e.g.*, a physical (direct), or phenotypic or chemical effect (indirect), such as the ability to increase or decrease cellular proliferation. It includes cell cycle arrest, the ability of cells to proliferate, and other characteristics of proliferating cells. “Functional effects” include *in vitro*, *in vivo*, and *ex vivo* activities.

[151] By “determining the functional effect” is meant assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of a colorectal cancer protein or nucleic acid, *e.g.*, physical, phenotypic and chemical effects. Such functional effects can be measured by any means known to those

skilled in the art, *e.g.*, physical effects such as changes in spectroscopic characteristics (*e.g.*, fluorescence, absorbance, refractive index); hydrodynamic (*e.g.*, shape); chromatographic; or solubility properties for the protein; measuring ligand binding activity or binding assays, *e.g.* binding to antibodies; measuring changes in ligand binding activity; and chemical or  
5 phenotypic effects such as measuring inducible markers or transcriptional activation of the protein; measuring cellular proliferation; measuring cell surface marker expression; measurement of changes in protein levels for colorectal cancer-associated sequences; measurement of RNA stability; phosphorylation or dephosphorylation; signal transduction, *e.g.*, receptor-ligand interactions, second messenger concentrations (*e.g.*, cAMP, IP3, or  
10 intracellular  $\text{Ca}^{2+}$ ); identification of downstream or reporter gene expression (CAT, luciferase,  $\beta$ -gal, GFP and the like), *e.g.*, via chemiluminescence, fluorescence, colorimetric reactions, antibody binding, and inducible markers.

[152] “Inhibitors”, “activators”, and “modulators” of colorectal cancer polynucleotide and polypeptide sequences are used to refer to activating, inhibitory, or  
15 modulating molecules identified using *in vitro* and *in vivo* assays of colorectal cancer polynucleotide and polypeptide sequences. Inhibitors are compounds that, *e.g.*, bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of colorectal cancer proteins or nucleic acids, *e.g.*, antagonists. “Activators” are compounds that increase, open, activate, facilitate, enhance  
20 activation, sensitize, agonize, or up regulate colorectal cancer protein or nucleic acid activity. Inhibitors, activators, or modulators also include genetically modified versions of colorectal cancer proteins, *e.g.*, versions with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, antibodies, antisense molecules, peptides, ribozymes, small chemical molecules and the like. Such assays for inhibitors and activators include, *e.g.*,  
25 expressing colorectal cancer protein *in vitro*, in cells, or cell membranes, applying putative modulator compounds, and then determining the functional effects on activity, as described above.

[153] Samples or assays comprising colorectal cancer proteins or colorectal cancer nucleic acids that are treated with a potential activator, inhibitor, or modulator are  
30 compared to control samples without the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative activity value of 100%. Inhibition of colorectal cancer is achieved when the activity value relative to the control is about 80%, preferably 50%, more preferably 25-0%. Activation of

colorectal cancer is achieved when the activity value relative to the control (untreated with activators) is 110%, more preferably 150%, more preferably 200-500% (i.e., two to five fold higher relative to the control), more preferably 1000-3000% higher.

[154] As will be appreciated by those in the art, this may be done by evaluation at either the gene or the protein level; that is, the amount of gene expression may be monitored using nucleic acid probes and the quantification of gene expression levels, or, alternatively, the gene product itself can be monitored, for example through the use of antibodies to the colorectal cancer protein and standard immunoassays.

[155] In a preferred embodiment, gene expression monitoring is done and a number of genes, i.e. an expression profile, is monitored simultaneously, although multiple protein expression monitoring can be done as well.

[156] In this embodiment, the colorectal cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of colorectal cancer sequences in a particular cell. The assays are further described below.

[157] Generally, in a preferred embodiment, a candidate bioactive agent is added to the cells prior to analysis. Moreover, screens are provided to identify a candidate bioactive agent which modulates colorectal cancer, modulates colorectal cancer proteins, binds to a colorectal cancer protein, or interferes between the binding of a colorectal cancer protein and an antibody.

[158] The term "candidate bioactive agent" or "test compound" or "drug candidate" or "modulator" or grammatical equivalents as used herein describes any molecule, either naturally occurring or synthetic, e.g., protein, oligopeptide (e.g., from about 5 to about 25 amino acids in length, preferably from about 10 to 20 or 12 to 18 amino acids in length, preferably 12, 15, or 18 amino acids in length), small organic molecule, polysaccharide, lipid, fatty acid, polynucleotide, oligonucleotide, etc., to be tested for the capacity to directly or indirectly modulate colorectal cancer sequences, including both nucleic acid and protein sequences. The test compound can be in the form of a library of test compounds, such as a combinatorial or randomized library that provides a sufficient range of diversity. Test compounds are optionally linked to a fusion partner, e.g., targeting compounds, rescue compounds, dimerization compounds, stabilizing compounds, addressable compounds, and other functional moieties. Conventionally, new chemical entities with useful properties are generated by identifying a test compound (called a "lead compound") with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and

evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

[159] In preferred embodiments, the bioactive agents modulate the expression profiles, or expression profile nucleic acids or proteins provided herein. In a particularly preferred embodiment, the candidate agent suppresses a colorectal cancer phenotype, for example to a normal colon tissue fingerprint. Similarly, the candidate agent preferably suppresses a severe colorectal cancer phenotype. Generally a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration or below the level of detection.

[160] In one aspect, a candidate agent will neutralize the effect of a colorectal cancer protein. By "neutralize" is meant that activity of a protein is either inhibited or counter acted against so as to have substantially no effect on a cell.

[161] Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 100 and less than about 2,500 daltons. Preferred small molecules are less than 2000, or less than 1500 or less than 1000 or less than 500 D. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof. Particularly preferred are peptides.

[162] Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification to produce structural analogs.

[163] In a preferred embodiment, the candidate bioactive agents are proteins. By "protein" herein is meant at least two covalently attached amino acids, which includes proteins, polypeptides, oligopeptides and peptides. The protein may be made up of naturally occurring amino acids and peptide bonds, or synthetic peptidomimetic structures.

5 Thus "amino acid", or "peptide residue", as used herein means both naturally occurring and synthetic amino acids. For example, homo-phenylalanine, citrulline and noreleucine are considered amino acids for the purposes of the invention. "Amino acid" also includes imino acid residues such as proline and hydroxyproline. The side chains may be in either the (R) or the (S) configuration. In the preferred embodiment, the amino acids are in the (S) or L-  
10 configuration. If non-naturally occurring side chains are used, non-amino acid substituents may be used, for example to prevent or retard in vivo degradations.

[164] In a preferred embodiment, the candidate bioactive agents are naturally occurring proteins or fragments of naturally occurring proteins. Thus, for example, cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts,  
15 may be used. In this way libraries of procaryotic and eucaryotic proteins may be made for screening in the methods of the invention. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred.

[165] In a preferred embodiment, the candidate bioactive agents are peptides  
20 of from about 5 to about 30 amino acids, with from about 5 to about 20 amino acids being preferred, and from about 7 to about 15 being particularly preferred. The peptides may be digests of naturally occurring proteins as is outlined above, random peptides, or "biased" random peptides. By "randomized" or grammatical equivalents herein is meant that each nucleic acid and peptide consists of essentially random nucleotides and amino acids,  
25 respectively. Since generally these random peptides (or nucleic acids, discussed below) are chemically synthesized, they may incorporate any nucleotide or amino acid at any position. The synthetic process can be designed to generate randomized proteins or nucleic acids, to allow the formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized candidate bioactive proteinaceous agents.

30 [166] In one embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a limited number of possibilities. For example, in a preferred embodiment, the nucleotides or amino acid residues are randomized within a defined class, for example, of hydrophobic

amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of nucleic acid binding domains, the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc., or to purines, etc.

5           [167] In a preferred embodiment, the candidate bioactive agents are nucleic acids, as defined above.

          [168] As described above generally for proteins, nucleic acid candidate bioactive agents may be naturally occurring nucleic acids, random nucleic acids, or “biased” random nucleic acids. For example, digests of procaryotic or eucaryotic genomes may be  
10       used as is outlined above for proteins.

          [169] In a preferred embodiment, the candidate bioactive agents are organic chemical moieties, a wide variety of which are available in the literature.

          [170] “Antibody” refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an  
15       antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively. Typically, the antigen-binding region of  
20       an antibody will be most critical in specificity and affinity of binding.

          [171] An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one “light” (about 25 kD) and one “heavy” chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily  
25       responsible for antigen recognition. The terms variable light chain ( $V_L$ ) and variable heavy chain ( $V_H$ ) refer to these light and heavy chains respectively.

          [172] Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to  
30       produce  $F(ab)'_2$ , a dimer of Fab which itself is a light chain joined to  $V_H$ - $C_H1$  by a disulfide bond. The  $F(ab)'_2$  may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the  $F(ab)'_2$  dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (*see Fundamental Immunology* (Paul ed., 3d ed. 1993). While various antibody fragments are defined in terms of the

digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized *de novo* using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (see, e.g., McCafferty *et al.*, *Nature* 348:552-554 (1990))

[173] For preparation of antibodies, e.g., recombinant, monoclonal, or polyclonal antibodies, many technique known in the art can be used (see, e.g., Kohler & Milstein, *Nature* 256:495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, pp. 77-96 in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985); Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *Antibodies, A Laboratory Manual* (1988); and Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986)). The genes encoding the heavy and light chains of an antibody of interest can be cloned from a cell, e.g., the genes encoding a monoclonal antibody can be cloned from a hybridoma and used to produce a recombinant monoclonal antibody. Gene libraries encoding heavy and light chains of monoclonal antibodies can also be made from hybridoma or plasma cells. Random combinations of the heavy and light chain gene products generate a large pool of antibodies with different antigenic specificity (see, e.g., Kubly, *Immunology* (3<sup>rd</sup> ed. 1997)). Techniques for the production of single chain antibodies or recombinant antibodies (U.S. Patent 4,946,778, U.S. Patent No. 4,816,567) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized or human antibodies (see, e.g., U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, Marks *et al.*, *Bio/Technology* 10:779-783 (1992); Lonberg *et al.*, *Nature* 368:856-859 (1994); Morrison, *Nature* 368:812-13 (1994); Fishwild *et al.*, *Nature Biotechnology* 14:845-51 (1996); Neuberger, *Nature Biotechnology* 14:826 (1996); and Lonberg & Huszar, *Intern. Rev. Immunol.* 13:65-93 (1995)). Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (see, e.g., McCafferty *et al.*, *Nature* 348:552-554 (1990); Marks *et al.*, *Biotechnology* 10:779-783 (1992)). Antibodies can also be made bispecific, i.e., able to recognize two different antigens (see, e.g., WO 93/08829, Traunecker *et al.*, *EMBO J.* 10:3655-3659 (1991); and Suresh *et al.*, *Methods in Enzymology* 121:210 (1986)). Antibodies can also be heteroconjugates, e.g., two covalently joined antibodies, or immunotoxins (see, e.g., U.S. Patent No. 4,676,980, WO 91/00360; WO 92/200373; and EP 03089).

[174] Methods for humanizing or primatizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as import residues, which are typically taken from an import variable domain. Humanization can be essentially performed following the method of Winter and co-workers (*see, e.g., Jones et al., Nature* 321:522-525 (1986); Riechmann *et al., Nature* 332:323-327 (1988); Verhoeyen *et al., Science* 239:1534-1536 (1988) and Presta, *Curr. Op. Struct. Biol.* 2:593-596 (1992)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such humanized antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

[175] A "chimeric antibody" is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, e.g., an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

[176] In one embodiment, the antibody is conjugated to an "effector" moiety. The effector moiety can be any number of molecules, including labeling moieties such as radioactive labels or fluorescent labels, or can be a therapeutic moiety. In one aspect the antibody modulates the activity of the protein.

[177] After the candidate agent has been added and the cells allowed to incubate for some period of time, the sample containing the target sequences to be analyzed is added to the biochip. If required, the target sequence is prepared using known techniques. For example, the sample may be treated to lyse the cells, using known lysis buffers, electroporation, etc., with purification and/or amplification such as PCR occurring as needed, as will be appreciated by those in the art. For example, an in vitro transcription with labels covalently attached to the nucleosides is done. Generally, the nucleic acids are labeled with biotin-FITC or PE, or with cy3 or cy5.

[178] In a preferred embodiment, the target sequence is labeled with, for example, a fluorescent, a chemiluminescent, a chemical, or a radioactive signal, to provide a means of detecting the target sequence's specific binding to a probe. The label also can be an enzyme, such as, alkaline phosphatase or horseradish peroxidase, which when provided with an appropriate substrate produces a product that can be detected. Alternatively, the label can be a labeled compound or small molecule, such as an enzyme inhibitor, that binds but is not catalyzed or altered by the enzyme. The label also can be a moiety or compound, such as, an epitope tag or biotin which specifically binds to streptavidin. For the example of biotin, the streptavidin is labeled as described above, thereby, providing a detectable signal for the bound target sequence. As known in the art, unbound labeled streptavidin is removed prior to analysis.

[179] As will be appreciated by those in the art, these assays can be direct hybridization assays or can comprise "sandwich assays", which include the use of multiple probes, as is generally outlined in U.S. Patent Nos. 5,681,702, 5,597,909, 5,545,730, 5,594,117, 5,591,584, 5,571,670, 5,580,731, 5,571,670, 5,591,584, 5,624,802, 5,635,352, 5,594,118, 5,359,100, 5,124,246 and 5,681,697, all of which are hereby incorporated by reference. In this embodiment, in general, the target nucleic acid is prepared as outlined above, and then added to the biochip comprising a plurality of nucleic acid probes, under conditions that allow the formation of a hybridization complex.

[180] A variety of hybridization conditions may be used in the present invention, including high, moderate and low stringency conditions as outlined above. The assays are generally run under stringency conditions which allows formation of the label probe hybridization complex only in the presence of target. Stringency can be controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration pH, organic solvent concentration, etc.

[181] These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Patent No. 5,681,697. Thus it may be desirable to perform certain steps at higher stringency conditions to reduce non-specific binding.

[182] The reactions outlined herein may be accomplished in a variety of ways, as will be appreciated by those in the art. Components of the reaction may be added simultaneously, or sequentially, in any order, with preferred embodiments outlined below. In addition, the reaction may include a variety of other reagents may be included in the assays. These include reagents like salts, buffers, neutral proteins, e.g. albumin, detergents, etc which

may be used to facilitate optimal hybridization and detection, and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used, depending on the sample preparation methods and purity of the target.

5           [183] Once the assay is run, the data is analyzed to determine the expression levels, and changes in expression levels as between states, of individual genes, forming a gene expression profile.

          [184] The screens are done to identify drugs or bioactive agents that modulate the colorectal cancer phenotype. Specifically, there are several types of screens  
10 that can be run. A preferred embodiment is in the screening of candidate agents that can induce or suppress a particular expression profile, thus preferably generating the associated phenotype. That is, candidate agents that can mimic or produce an expression profile in colorectal cancer similar to the expression profile of normal colon tissue is expected to result in a suppression of the colorectal cancer phenotype. Thus, in this embodiment, mimicking an  
15 expression profile, or changing one profile to another, is the goal.

          [185] In a preferred embodiment, as for the diagnosis and prognosis applications, having identified the differentially expressed genes important in any one state, screens can be run to alter the expression of the genes individually. That is, screening for modulation of regulation of expression of a single gene can be done; that is, rather than try to  
20 mimic all or part of an expression profile, screening for regulation of individual genes can be done. Thus, for example, particularly in the case of target genes whose presence or absence is unique between two states, screening is done for modulators of the target gene expression.

          [186] In a preferred embodiment, screening is done to alter the biological function of the expression product of the differentially expressed gene. Again, having  
25 identified the importance of a gene in a particular state, screening for agents that bind and/or modulate the biological activity of the gene product can be run as is more fully outlined below.

          [187] Thus, screening of candidate agents that modulate the colorectal cancer phenotype either at the gene expression level or the protein level can be done.

30           [188] In addition screens can be done for novel genes that are induced in response to a candidate agent. After identifying a candidate agent based upon its ability to suppress a colorectal cancer expression pattern leading to a normal expression pattern, or modulate a single colorectal cancer gene expression profile so as to mimic the expression of the gene from normal tissue, a screen as described above can be performed to identify genes

that are specifically modulated in response to the agent. Comparing expression profiles between normal tissue and agent treated colorectal cancer tissue reveals genes that are not expressed in normal tissue or colorectal cancer tissue, but are expressed in agent treated tissue. These agent specific sequences can be identified and used by any of the methods  
5 described herein for colorectal cancer genes or proteins. In particular these sequences and the proteins they encode find use in marking or identifying agent treated cells. In addition, antibodies can be raised against the agent induced proteins and used to target novel therapeutics to the treated colorectal cancer tissue sample.

[189] Thus, in one embodiment, a candidate agent is administered to a  
10 population of colorectal cancer cells, that thus has an associated colorectal cancer expression profile. By “administration” or “contacting” herein is meant that the candidate agent is added to the cells in such a manner as to allow the agent to act upon the cell, whether by uptake and intracellular action, or by action at the cell surface. In some embodiments, nucleic acid encoding a proteinaceous candidate agent (i.e. a peptide) may be put into a viral  
15 construct such as a retroviral construct and added to the cell, such that expression of the peptide agent is accomplished; see PCT US97/01019, hereby expressly incorporated by reference.

[190] Once the candidate agent has been administered to the cells, the cells can be washed if desired and are allowed to incubate under preferably physiological  
20 conditions for some period of time. The cells are then harvested and a new gene expression profile is generated, as outlined herein.

[191] Thus, for example, colorectal cancer tissue may be screened for agents that reduce or suppress the colorectal cancer phenotype. A change in at least one gene of the expression profile indicates that the agent has an effect on colorectal cancer  
25 activity. By defining such a signature for the colorectal cancer phenotype, screens for new drugs that alter the phenotype can be devised. With this approach, the drug target need not be known and need not be represented in the original expression screening platform, nor does the level of transcript for the target protein need to change.

[192] In a preferred embodiment, as outlined above, screens may be done on  
30 individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of either the expression of the gene or the gene product itself can be done. The gene products of differentially expressed genes are sometimes referred to herein as “colorectal cancer modulator proteins”. The colorectal cancer modulator protein may be a fragment, or

alternatively, be the full length protein to a fragment shown herein. Preferably, the colorectal cancer modulator protein is a fragment of approximately 14 to 24 amino acids long. More preferably the fragment is a soluble fragment.

[193] In a preferred embodiment, the fragment is charged and from the c-terminus. In one embodiment, the c-terminus of the fragment is kept as a free acid and the n-terminus is a free amine to aid in coupling, i.e., to cysteine. In another embodiment, the fragment is an internal peptide overlapping hydrophilic stretch the protein. In a preferred embodiment, the termini is blocked. In another preferred embodiment, the fragment is a novel fragment from the N-terminal. In one embodiment, the fragment excludes sequence outside of the N-terminal, in another embodiment, the fragment includes at least a portion of the N-terminal. "N-terminal" is used interchangeably herein with "N-terminus" which is further described above.

[194] In one embodiment the colorectal cancer proteins are conjugated to an immunogenic agent as discussed herein. In one embodiment the colorectal cancer protein is conjugated to BSA.

[195] Thus, in a preferred embodiment, screening for modulators of expression of specific genes can be done. This will be done as outlined above, but in general the expression of only one or a few genes are evaluated.

[196] In a preferred embodiment, screens are designed to first find candidate agents that can bind to differentially expressed proteins, and then these agents may be used in assays that evaluate the ability of the candidate agent to modulate differentially expressed activity. Thus, as will be appreciated by those in the art, there are a number of different assays which may be run; binding assays and activity assays.

[197] In a preferred embodiment, binding assays are done. In general, purified or isolated gene product is used; that is, the gene products of one or more differentially expressed nucleic acids are made. In general, this is done as is known in the art. For example, antibodies are generated to the protein gene products, and standard immunoassays are run to determine the amount of protein present. Alternatively, cells comprising the colorectal cancer proteins can be used in the assays.

[198] Thus, in a preferred embodiment, the methods comprise combining a colorectal cancer protein and a candidate bioactive agent, and determining the binding of the candidate agent to the colorectal cancer protein. Preferred embodiments utilize the human colorectal cancer protein, although other mammalian proteins may also be used, for example

for the development of animal models of human disease. In some embodiments, as outlined herein, variant or derivative colorectal cancer proteins may be used.

[199] Generally, in a preferred embodiment of the methods herein, the colorectal cancer protein or the candidate agent is non-diffusably bound to an insoluble support having isolated sample receiving areas (e.g. a microtiter plate, an array, etc.). The insoluble supports may be made of any composition to which the compositions can be bound, is readily separated from soluble material, and is otherwise compatible with the overall method of screening. The surface of such supports may be solid or porous and of any convenient shape. Examples of suitable insoluble supports include microtiter plates, arrays, membranes and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharides, nylon or nitrocellulose, teflon, etc. Microtiter plates and arrays are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents and samples. The particular manner of binding of the composition is not crucial so long as it is compatible with the reagents and overall methods of the invention, maintains the activity of the composition and is nondiffusable. Preferred methods of binding include the use of antibodies (which do not sterically block either the ligand binding site or activation sequence when the protein is bound to the support), direct binding to "sticky" or ionic supports, chemical crosslinking, the synthesis of the protein or agent on the surface, etc. Following binding of the protein or agent, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein or other innocuous protein or other moiety.

[200] In a preferred embodiment, the colorectal cancer protein is bound to the support, and a candidate bioactive agent is added to the assay. Alternatively, the candidate agent is bound to the support and the colorectal cancer protein is added. Novel binding agents include specific antibodies, non-natural binding agents identified in screens of chemical libraries, peptide analogs, etc. Of particular interest are screening assays for agents that have a low toxicity for human cells. A wide variety of assays may be used for this purpose, including labeled in vitro protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.) and the like.

[201] The determination of the binding of the candidate bioactive agent to the colorectal cancer protein may be done in a number of ways. In a preferred embodiment, the candidate bioactive agent is labeled, and binding determined directly. For example, this may be done by attaching all or a portion of the colorectal cancer protein to a solid support,

adding a labeled candidate agent (for example a fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps may be utilized as is known in the art.

[202] By "labeled" herein is meant that the compound is either directly or indirectly labeled with a label which provides a detectable signal, e.g. radioisotope, fluorescers, enzyme, antibodies, particles such as magnetic particles, chemiluminescers, or specific binding molecules, etc. Specific binding molecules include pairs, such as biotin and streptavidin, digoxin and antidigoxin etc. For the specific binding members, the complementary member would normally be labeled with a molecule which provides for detection, in accordance with known procedures, as outlined above. The label can directly or indirectly provide a detectable signal.

[203] In some embodiments, only one of the components is labeled. For example, the proteins (or proteinaceous candidate agents) may be labeled at tyrosine positions using  $^{125}\text{I}$ , or with fluorophores. Alternatively, more than one component may be labeled with different labels; using  $^{125}\text{I}$  for the proteins, for example, and a fluorophor for the candidate agents.

[204] In a preferred embodiment, the binding of the candidate bioactive agent is determined through the use of competitive binding assays. In this embodiment, the competitor is a binding moiety known to bind to the target molecule (i.e. colorectal cancer), such as an antibody, peptide, binding partner, ligand, etc. Under certain circumstances, there may be competitive binding as between the bioactive agent and the binding moiety, with the binding moiety displacing the bioactive agent.

[205] In one embodiment, the candidate bioactive agent is labeled. Either the candidate bioactive agent, or the competitor, or both, is added first to the protein for a time sufficient to allow binding, if present. Incubations may be performed at any temperature which facilitates optimal activity, typically between 4 and 40°C. Incubation periods are selected for optimum activity, but may also be optimized to facilitate rapid high throughput screening. Typically between 0.1 and 1 hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

[206] In a preferred embodiment, the competitor is added first, followed by the candidate bioactive agent. Displacement of the competitor is an indication that the candidate bioactive agent is binding to the colorectal cancer protein and thus is capable of binding to, and potentially modulating, the activity of the colorectal cancer protein. In this

embodiment, either component can be labeled. Thus, for example, if the competitor is labeled, the presence of label in the wash solution indicates displacement by the agent. Alternatively, if the candidate bioactive agent is labeled, the presence of the label on the support indicates displacement.

5           [207] In an alternative embodiment, the candidate bioactive agent is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor may indicate that the bioactive agent is bound to the colorectal cancer protein with a higher affinity. Thus, if the candidate bioactive agent is labeled, the presence of the label on the support, coupled with a lack of competitor binding, may indicate that the  
10 candidate agent is capable of binding to the colorectal cancer protein.

          [208] In a preferred embodiment, the methods comprise differential screening to identity bioactive agents that are capable of modulating the activity of the colorectal cancer proteins. In this embodiment, the methods comprise combining a colorectal cancer protein and a competitor in a first sample. A second sample comprises a  
15 candidate bioactive agent, a colorectal cancer protein and a competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the colorectal cancer protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the agent is capable  
20 of binding to the colorectal cancer protein.

          [209] Alternatively, a preferred embodiment utilizes differential screening to identify drug candidates that bind to the native colorectal cancer protein, but cannot bind to modified colorectal cancer proteins. The structure of the colorectal cancer protein may be modeled, and used in rational drug design to synthesize agents that interact with that site.  
25 Drug candidates that affect colorectal cancer bioactivity are also identified by screening drugs for the ability to either enhance or reduce the activity of the protein.

          [210] Positive controls and negative controls may be used in the assays. Preferably all control and test samples are performed in at least triplicate to obtain statistically significant results. Incubation of all samples is for a time sufficient for the  
30 binding of the agent to the protein. Following incubation, all samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a radiolabel is employed, the samples may be counted in a scintillation counter to determine the amount of bound compound.

[211] A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as  
5 protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in any order that provides for the requisite binding.

[212] Screening for agents that modulate the activity of colorectal cancer proteins may also be done. In a preferred embodiment, methods for screening for a bioactive agent capable of modulating the activity of colorectal cancer proteins comprise the steps of  
10 adding a candidate bioactive agent to a sample of colorectal cancer proteins, as above, and determining an alteration in the biological activity of colorectal cancer proteins.

"Modulating the activity of colorectal cancer " includes an increase in activity, a decrease in activity, or a change in the type or kind of activity present. Thus, in this embodiment, the candidate agent should both bind to colorectal cancer proteins (although this may not be  
15 necessary), and alter its biological or biochemical activity as defined herein. The methods include both in vitro screening methods, as are generally outlined above, and in vivo screening of cells for alterations in the presence, distribution, activity or amount of colorectal cancer proteins.

[213] Thus, in this embodiment, the methods comprise combining a  
20 colorectal cancer sample and a candidate bioactive agent, and evaluating the effect on colorectal cancer activity. By "colorectal cancer activity" or grammatical equivalents herein is meant one of the colorectal cancer 's biological activities, including, but not limited to, cell division, preferably in colon tissue, cell proliferation, tumor growth, transformation of cells. In one embodiment, colorectal cancer activity includes activation of a gene identified by a  
25 nucleic acid of Table 1. An inhibitor of colorectal cancer activity is the inhibition of any one or more colorectal cancer activities.

[214] In a preferred embodiment, the activity of the colorectal cancer protein is increased; in another preferred embodiment, the activity of the colorectal cancer protein is decreased. Thus, bioactive agents that are antagonists are preferred in some embodiments,  
30 and bioactive agents that are agonists may be preferred in other embodiments.

[215] In a preferred embodiment, the invention provides methods for screening for bioactive agents capable of modulating the activity of a colorectal cancer protein. The methods comprise adding a candidate bioactive agent, as defined above, to a cell comprising colorectal cancer proteins. Preferred cell types include almost any cell. The

cells contain a recombinant nucleic acid that encodes a colorectal cancer protein. In a preferred embodiment, a library of candidate agents are tested on a plurality of cells.

[216] In one aspect, the assays are evaluated in the presence or absence or previous or subsequent exposure of physiological signals, for example hormones, antibodies, peptides, antigens, cytokines, growth factors, action potentials, pharmacological agents including chemotherapeutics, radiation, carcinogenics, or other cells (i.e. cell-cell contacts). In another example, the determinations are determined at different stages of the cell cycle process.

[217] In this way, bioactive agents are identified. Compounds with pharmacological activity are able to enhance or interfere with the activity of the colorectal cancer protein. In one embodiment, "colorectal cancer protein activity" as used herein includes at least one of the following: colorectal cancer activity, binding to the colorectal cancer protein, activation of the colorectal cancer protein or activation of substrates of the colorectal cancer protein by the colorectal cancer protein. In one embodiment, colorectal cancer activity is defined as the unregulated proliferation of colon tissue, or the growth of cancer in colon tissue. In one aspect, colorectal cancer activity as defined herein is related to the activity of the colorectal cancer protein in the upregulation of the colorectal cancer protein in colon cancer tissue.

[218] In another embodiment, colorectal cancer protein activity includes at least one of the following: colorectal cancer activity, binding to the CBF9 nucleic acid or poly peptide of Table 2 or binding to a nucleic acid of Table 1, or a peptide encoded by a nucleic acid of Table 1 or activation of substrates of the gene products identified by a nucleic acid of Table 1 or substrates of CBF9, which is shown in Table 2. In one aspect, colorectal cancer activity as defined herein is related to the activity of genes defined by the nucleic acids of Table 1 or of CBF9 as defined in Table 2, in colon cancer tissue.

[219] In one embodiment, a method of inhibiting colon cancer cell division is provided. The method comprises administration of a colorectal cancer inhibitor.

[220] In another embodiment, a method of inhibiting tumor growth is provided. The method comprises administration of a colorectal cancer inhibitor.

[221] In a further embodiment, methods of treating cells or individuals with cancer are provided. The method comprises administration of a colorectal cancer inhibitor.

[222] In one embodiment, a colorectal cancer inhibitor is an antibody as discussed above. In another embodiment, the colorectal cancer inhibitor is an antisense molecule. Antisense molecules as used herein include antisense or sense oligonucleotides

comprising a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target mRNA (sense) or DNA (antisense) sequences for colorectal cancer molecules. A preferred antisense molecule is for the colorectal cancer sequences referenced in Table 1 or Table 2, or for a ligand or activator thereof. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment generally at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (Cancer Res. 48:2659, 1988) and van der Krol et al. (BioTechniques 6:958, 1988).

[223] Antisense molecules may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell. Alternatively, a sense or an antisense oligonucleotide may be introduced into a cell containing the target nucleic acid sequence by formation of an oligonucleotide-lipid complex, as described in WO 90/10448. It is understood that the use of antisense molecules or knock out and knock in models may also be used in screening assays as discussed above, in addition to methods of treatment.

[224] The compounds having the desired pharmacological activity may be administered in a physiologically acceptable carrier to a host, as previously described. The agents may be administered in a variety of ways, orally, parenterally e.g., subcutaneously, intraperitoneally, intravascularly, etc. Depending upon the manner of introduction, the compounds may be formulated in a variety of ways. The concentration of therapeutically active compound in the formulation may vary from about 0.1-100 wt.%. The agents may be administered alone or in combination with other treatments, i.e., radiation.

[225] The pharmaceutical compositions can be prepared in various forms, such as granules, tablets, pills, suppositories, capsules, suspensions, salves, lotions and the like. Pharmaceutical grade organic or inorganic carriers and/or diluents suitable for oral and topical use can be used to make up compositions containing the therapeutically-active compounds. Diluents known to the art include aqueous media, vegetable and animal oils and fats. Stabilizing agents, wetting and emulsifying agents, salts for varying the osmotic

pressure or buffers for securing an adequate pH value, and skin penetration enhancers can be used as auxiliary agents.

[226] Without being bound by theory, it appears that the various colorectal cancer sequences are important in colorectal cancer. Accordingly, disorders based on mutant or variant colorectal cancer genes may be determined. In one embodiment, the invention provides methods for identifying cells containing variant colorectal cancer genes comprising determining all or part of the sequence of at least one endogenous colorectal cancer genes in a cell. As will be appreciated by those in the art, this may be done using any number of sequencing techniques. In a preferred embodiment, the invention provides methods of identifying the colorectal cancer genotype of an individual comprising determining all or part of the sequence of at least one colorectal cancer gene of the individual. This is generally done in at least one tissue of the individual, and may include the evaluation of a number of tissues or different samples of the same tissue. The method may include comparing the sequence of the sequenced colorectal cancer gene to a known colorectal cancer gene, i.e. a wild-type gene.

[227] The sequence of all or part of the colorectal cancer gene can then be compared to the sequence of a known colorectal cancer gene to determine if any differences exist. This can be done using any number of known homology programs, such as Bestfit, etc. In a preferred embodiment, the presence of a difference in the sequence between the colorectal cancer gene of the patient and the known colorectal cancer gene is indicative of a disease state or a propensity for a disease state, as outlined herein.

[228] In a preferred embodiment, the colorectal cancer genes are used as probes to determine the number of copies of the colorectal cancer gene in the genome.

[229] In another preferred embodiment colorectal cancer genes are used as probed to determine the chromosomal localization of the colorectal cancer genes. Information such as chromosomal localization finds use in providing a diagnosis or prognosis in particular when chromosomal abnormalities such as translocations, and the like are identified in colorectal cancer gene loci.

[230] Thus, in one embodiment, methods of modulating colorectal cancer in cells or organisms are provided. In one embodiment, the methods comprise administering to a cell an anti-colorectal cancer antibody that reduces or eliminates the biological activity of an endogenous colorectal cancer protein. Alternatively, the methods comprise administering to a cell or organism a recombinant nucleic acid encoding a colorectal cancer protein. As will be appreciated by those in the art, this may be accomplished in any number

of ways. In a preferred embodiment, for example when the colorectal cancer sequence is down-regulated in colorectal cancer, the activity of the colorectal cancer gene is increased by increasing the amount of colorectal cancer in the cell, for example by overexpressing the endogenous colorectal cancer or by administering a gene encoding the colorectal cancer sequence, using known gene-therapy techniques, for example. In a preferred embodiment, the gene therapy techniques include the incorporation of the exogenous gene using enhanced homologous recombination (EHR), for example as described in PCT/US93/03868, hereby incorporated by reference in its entirety. Alternatively, for example when the colorectal cancer sequence is up-regulated in colorectal cancer, the activity of the endogenous colorectal cancer gene is decreased, for example by the administration of a colorectal cancer antisense nucleic acid.

[231] In one embodiment, the colorectal cancer proteins of the present invention may be used to generate polyclonal and monoclonal antibodies to colorectal cancer proteins, which are useful as described herein. Similarly, the colorectal cancer proteins can be coupled, using standard technology, to affinity chromatography columns. These columns may then be used to purify colorectal cancer antibodies. In a preferred embodiment, the antibodies are generated to epitopes unique to a colorectal cancer protein; that is, the antibodies show little or no cross-reactivity to other proteins. These antibodies find use in a number of applications. For example, the colorectal cancer antibodies may be coupled to standard affinity chromatography columns and used to purify colorectal cancer proteins. The antibodies may also be used as blocking polypeptides, as outlined above, since they will specifically bind to the colorectal cancer protein.

[232] In one embodiment, a therapeutically effective dose of a colorectal cancer or modulator thereof is administered to a patient. By "therapeutically effective dose" herein is meant a dose that produces the effects for which it is administered. The exact dose will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using known techniques. As is known in the art, adjustments for colorectal cancer degradation, systemic versus localized delivery, and rate of new protease synthesis, as well as the age, body weight, general health, sex, diet, time of administration, drug interaction and the severity of the condition may be necessary, and will be ascertainable with routine experimentation by those skilled in the art.

[233] A "patient" for the purposes of the present invention includes both humans and other animals, particularly mammals, and organisms. Thus the methods are

applicable to both human therapy and veterinary applications. In the preferred embodiment the patient is a mammal, and in the most preferred embodiment the patient is human.

[234] The administration of the colorectal cancer proteins and modulators of the present invention can be done in a variety of ways as discussed above, including, but not limited to, orally, subcutaneously, intravenously, intranasally, transdermally, intraperitoneally, intramuscularly, intrapulmonary, vaginally, rectally, or intraocularly. In some instances, for example, in the treatment of wounds and inflammation, the colorectal cancer proteins and modulators may be directly applied as a solution or spray.

[235] The pharmaceutical compositions of the present invention comprise a colorectal cancer protein in a form suitable for administration to a patient. In the preferred embodiment, the pharmaceutical compositions are in a water soluble form, such as being present as pharmaceutically acceptable salts, which is meant to include both acid and base addition salts. "Pharmaceutically acceptable acid addition salt" refers to those salts that retain the biological effectiveness of the free bases and that are not biologically or otherwise undesirable, formed with inorganic acids such as hydrochloric acid, hydrobromic acid, sulfuric acid, nitric acid, phosphoric acid and the like, and organic acids such as acetic acid, propionic acid, glycolic acid, pyruvic acid, oxalic acid, maleic acid, malonic acid, succinic acid, fumaric acid, tartaric acid, citric acid, benzoic acid, cinnamic acid, mandelic acid, methanesulfonic acid, ethanesulfonic acid, p-toluenesulfonic acid, salicylic acid and the like. "Pharmaceutically acceptable base addition salts" include those derived from inorganic bases such as sodium, potassium, lithium, ammonium, calcium, magnesium, iron, zinc, copper, manganese, aluminum salts and the like. Particularly preferred are the ammonium, potassium, sodium, calcium, and magnesium salts. Salts derived from pharmaceutically acceptable organic non-toxic bases include salts of primary, secondary, and tertiary amines, substituted amines including naturally occurring substituted amines, cyclic amines and basic ion exchange resins, such as isopropylamine, trimethylamine, diethylamine, triethylamine, tripropylamine, and ethanolamine.

[236] The pharmaceutical compositions may also include one or more of the following: carrier proteins such as serum albumin; buffers; fillers such as microcrystalline cellulose, lactose, corn and other starches; binding agents; sweeteners and other flavoring agents; coloring agents; and polyethylene glycol. Additives are well known in the art, and are used in a variety of formulations.

[237] In a preferred embodiment, colorectal cancer proteins and modulators are administered as therapeutic agents, and can be formulated as outlined above. Similarly,

colorectal cancer genes (including both the full-length sequence, partial sequences, or regulatory sequences of the colorectal cancer coding regions) can be administered in gene therapy applications, as is known in the art. These colorectal cancer genes can include antisense applications, either as gene therapy (i.e. for incorporation into the genome) or as antisense compositions, as will be appreciated by those in the art.

[238] In a preferred embodiment, colorectal cancer genes are administered as DNA vaccines, either single genes or combinations of colorectal cancer genes. Naked DNA vaccines are generally known in the art. Brower, Nature Biotechnology, 16:1304-1305 (1998).

[239] In one embodiment, colorectal cancer genes of the present invention are used as DNA vaccines. Methods for the use of genes as DNA vaccines are well known to one of ordinary skill in the art, and include placing a colorectal cancer gene or portion of a colorectal cancer gene under the control of a promoter for expression in a colorectal cancer patient. The colorectal cancer gene used for DNA vaccines can encode full-length colorectal cancer proteins, but more preferably encodes portions of the colorectal cancer proteins including peptides derived from the colorectal cancer protein. In a preferred embodiment a patient is immunized with a DNA vaccine comprising a plurality of nucleotide sequences derived from a colorectal cancer gene. Similarly, it is possible to immunize a patient with a plurality of colorectal cancer genes or portions thereof as defined herein. Without being bound by theory, expression of the polypeptide encoded by the DNA vaccine, cytotoxic T-cells, helper T-cells and antibodies are induced which recognize and destroy or eliminate cells expressing colorectal cancer proteins.

[240] In a preferred embodiment, the DNA vaccines include a gene encoding an adjuvant molecule with the DNA vaccine. Such adjuvant molecules include cytokines that increase the immunogenic response to the colorectal cancer polypeptide encoded by the DNA vaccine. Additional or alternative adjuvants are known to those of ordinary skill in the art and find use in the invention.

[241] In another preferred embodiment colorectal cancer genes find use in generating animal models of colorectal cancer. As is appreciated by one of ordinary skill in the art, when the colorectal cancer gene identified is repressed or diminished in colorectal cancer tissue, gene therapy technology wherein antisense RNA directed to the colorectal cancer gene will also diminish or repress expression of the gene. An animal generated as such serves as an animal model of colorectal cancer that finds use in screening bioactive drug candidates. Similarly, gene knockout technology, for example as a result of

homologous recombination with an appropriate gene targeting vector, will result in the absence of the colorectal cancer protein. When desired, tissue-specific expression or knockout of the colorectal cancer protein may be necessary.

[242] It is also possible that the colorectal cancer protein is overexpressed in colorectal cancer. As such, transgenic animals can be generated that overexpress the colorectal cancer protein. Depending on the desired expression level, promoters of various strengths can be employed to express the transgene. Also, the number of copies of the integrated transgene can be determined and compared for a determination of the expression level of the transgene. Animals generated by such methods find use as animal models of colorectal cancer and are additionally useful in screening for bioactive molecules to treat colorectal cancer.

### EXAMPLES

[243] It is understood that the examples described herein in no way serve to limit the true scope of this invention, but rather are presented for illustrative purposes. All references and sequences of accession numbers cited herein are incorporated by reference in their entirety.

#### [244] Example 1

##### Tissue Preparation, Labeling Chips, and Fingerprints

[245] Purify total RNA from tissue using TRIzol Reagent

[246] Estimate tissue weight. Homogenize tissue samples in 1ml of TRIzol per 50mg of tissue using a Polytron 3100 homogenizer. The generator/probe used depends upon the tissue size. A generator that is too large for the amount of tissue to be homogenized will cause a loss of sample and lower RNA yield. Use the 20mm generator for tissue weighing more than 0.6g. If the working volume is greater than 2ml, then homogenize tissue in a 15ml polypropylene tube (Falcon 2059). Fill tube no greater than 10ml.

### HOMOGENIZATION

[247] Before using generator, it should have been cleaned after last usage by running it through soapy H2O and rinsing thoroughly. Run through with EtOH to sterilize. Keep tissue frozen until ready. Add TRIzol directly to frozen tissue then homogenize.

[248] Following homogenization, remove insoluble material from the homogenate by centrifugation at 7500 x g for 15 min. in a Sorvall superspeed or 12,000 x g for 10 min. in an Eppendorf centrifuge at 4oC. Transfer the cleared homogenate to a new tube(s). The samples may be frozen now at -60 to -70oC (and kept for at least one month) or  
5 you may continue with the purification.

### **PHASE SEPARATION**

[249] Incubate the homogenized samples for 5 minutes at room temperature.

[250] Add 0.2ml of chloroform per 1ml of TRIzol reagent used in the  
10 original homogenization.

[251] Cap tubes securely and shake tubes vigorously by hand (do not vortex) for 15 seconds.

[252] Incubate samples at room temp. for 2-3 minutes. Centrifuge samples at 6500rpm in a Sorvall superspeed for 30 min. at 4oC. (You may spin at up to 12,000 x g  
15 for 10 min. but you risk breaking your tubes in the centrifuge.)

### **RNA PRECIPITATION**

[253] Transfer the aqueous phase to a fresh tube. Save the organic phase if isolation of DNA or protein is desired. Add 0.5ml of isopropyl alcohol per 1ml of TRIzol reagent used in the original homogenization. Cap tubes securely and invert to mix. Incubate  
20 samples at room temp. for 10 minutes. Centrifuge samples at 6500rpm in Sorvall for 20min. at 4oC.

### **RNA WASH**

[254] Pour off the supernate. Wash pellet with cold 75% ethanol. Use 1ml of 75% ethanol per 1ml of TRIzol reagent used in the initial homogenization. Cap tubes securely and invert several times to loosen pellet. (Do not vortex). Centrifuge at <8000rpm (<7500 x g) for 5 minutes at 4oC.  
25

[255] Pour off the wash. Carefully transfer pellet to an eppendorf tube (let it  
30 slide down the tube into the new tube and use a pipet tip to help guide it in if necessary). Depending on the volumes you are working with, you can decide what size tube(s) you want to precipitate the RNA in. When I tried leaving the RNA in the large 15ml tube, it took so long to dry (i.e. it did not dry) that I eventually had to transfer it to a smaller tube. Let pellet

dry in hood. Resuspend RNA in an appropriate volume of DEPC H<sub>2</sub>O. Try for 2-5ug/ul.  
Take absorbance readings.

[256] Purify poly A+ mRNA from total RNA or clean up total RNA with  
5 Qiagen' s RNeasy kit

[257] Purification of poly A+ mRNA from total RNA. Heat oligotex  
suspension to 37oC and mix immediately before adding to RNA. Incubate Elution Buffer at  
70oC. Warm up 2 x Binding Buffer at 65oC if there is precipitate in the buffer. Mix total  
RNA with DEPC-treated water, 2 x Binding Buffer, and Oligotex according to Table 2 on  
10 page 16 of the Oligotex Handbook. Incubate for 3 minutes at 65oC. Incubate for 10 minutes  
at room temperature.

[258] Centrifuge for 2 minutes at 14,000 to 18,000 g. If centrifuge has a  
"soft setting," then use it. Remove supernatant without disturbing Oligotex pellet. A little bit  
15 of solution can be left behind to reduce the loss of Oligotex. Save sup until certain that  
satisfactory binding and elution of poly A+ mRNA has occurred.

[259] Gently resuspend in Wash Buffer OW2 and pipet onto spin column.  
Centrifuge the spin column at full speed (soft setting if possible) for 1 minute.  
20

[260] Transfer spin column to a new collection tube and gently resuspend in  
Wash Buffer OW2 and centrifuge as describe herein.

[261] Transfer spin column to a new tube and elute with 20 to 100 ul of  
25 preheated (70oC) Elution Buffer. Gently resuspend Oligotex resin by pipetting up and down.  
Centrifuge as above. Repeat elution with fresh elution buffer or use first eluate to keep the  
elution volume low.

[262] Read absorbance, using diluted Elution Buffer as the blank.  
30

[263] Before proceeding with cDNA synthesis, the mRNA must be  
precipitated. Some component leftover or in the Elution Buffer from the Oligotex  
purification procedure will inhibit downstream enzymatic reactions of the mRNA.

**Ethanol Precipitation**

[264] Add 0.4 vol. of 7.5 M NH<sub>4</sub>OAc + 2.5 vol. of cold 100% ethanol. Precipitate at -20°C 1 hour to overnight (or 20-30 min. at -70°C). Centrifuge at 14,000-16,000 x g for 30 minutes at 4°C. Wash pellet with 0.5ml of 80% ethanol (-20°C) then  
 5 centrifuge at 14,000-16,000 x g for 5 minutes at room temperature. Repeat 80% ethanol wash. Dry the last bit of ethanol from the pellet in the hood. (Do not speed vacuum). Suspend pellet in DEPC H<sub>2</sub>O at 1ug/ul concentration.

Clean up total RNA using Qiagen's RNeasy kit

10 [265] Add no more than 100ug to an RNeasy column. Adjust sample to a volume of 100ul with RNase-free water. Add 350ul Buffer RLT then 250ul ethanol (100%) to the sample. Mix by pipetting (do not centrifuge) then apply sample to an RNeasy mini spin column. Centrifuge for 15 sec at >10,000rpm. If concerned about yield, re-apply flowthrough to column and centrifuge again.

15 [266] Transfer column to a new 2-ml collection tube. Add 500ul Buffer RPE and centrifuge for 15 sec at >10,000rpm. Discard flowthrough. Add 500ul Buffer RPE and centrifuge for 15 sec at >10,000rpm. Discard flowthrough then centrifuge for 2 min at maximum speed to dry column membrane. Transfer column to a new 1.5-ml collection tube and apply 30-50ul of RNase-free water directly onto column membrane. Centrifuge 1 min at  
 20 >10,000rpm. Repeat elution.

[267] Take absorbance reading. If necessary, ethanol precipitate with ammonium acetate and 2.5X volume 100% ethanol.

25 [268] Make cDNA using Gibco's "SuperScript Choice System for cDNA Synthesis" kit

**First Strand cDNA Synthesis**

[269] Use 5ug of total RNA or 1ug of polyA+ mRNA as starting material. For total RNA, use 2ul of SuperScript RT. For polyA+ mRNA, use 1ul of SuperScript RT. Final volume of first strand synthesis mix is 20ul. RNA must be in a volume no greater than  
 30 10ul. Incubate RNA with 1ul of 100pmol T7-T24 oligo for 10 min at 70°C. On ice, add 7 ul of: 4ul 5X 1st Strand Buffer, 2ul of 0.1M DTT, and 1 ul of 10mM dNTP mix. Incubate at 37°C for 2 min then add SuperScript RT

Incubate at 37°C for 1 hour.

Second Strand Synthesis

Place 1st strand reactions on ice.

Add: 91ul DEPC H<sub>2</sub>O

30ul 5X 2nd Strand Buffer

3ul 10mM dNTP mix

5 1ul 10U/ul E.coli DNA Ligase

4ul 10U/ul E.coli DNA Polymerase

1ul 2U/ul RNase H

10 [270] Make the above into a mix if there are more than 2 samples. Mix and incubate 2 hours at 16C.

[271] Add 2ul T4 DNA Polymerase. Incubate 5 min at 16C. Add 10ul of 0.5M EDTA

[272] Clean up cDNA

15 [273] Phenol:Chloroform:Isoamyl Alcohol (25:24:1) purification using Phase-Lock gel tubes:

[274] Centrifuge PLG tubes for 30 sec at maximum speed. Transfer cDNA mix to PLG tube. Add equal volume of phenol:chloroform:isamyl alcohol and shake vigorously (do not vortex). Centrifuge 5 minutes at maximum speed. Transfer top aqueous solution to a new tube. Ethanol precipitate: add 7.5X 5M NH<sub>4</sub>Oac and 2.5X volume of 100% ethanol. Centrifuge immediately at room temp. for 20 min, maximum speed. Remove sup then wash pellet 2X with cold 80% ethanol. Remove as much ethanol wash as possible then let pellet air dry. Resuspend pellet in 3ul RNase-free water.

25 In vitro Transcription (IVT) and labeling with biotin

Pipet 1.5ul of cDNA into a thin-wall PCR tube.

Make NTP labeling mix:

30 Combine at room temperature: 2ul T7 10xATP (75mM) (Ambion)  
2ul T7 10xGTP (75mM) (Ambion)  
1.5ul T7 10xCTP (75mM) (Ambion)  
1.5ul T7 10xUTP (75mM) (Ambion)  
3.75ul 10mM Bio-11-UTP (Boehringer-Mannheim/Roche or Enzo)  
3.75ul 10mM Bio-16-CTP (Enzo)

2ul 10x T7 transcription buffer (Ambion)  
2ul 10x T7 enzyme mix (Ambion)

5 [275] Final volume of total reaction is 20ul. Incubate 6 hours at 37C in a PCR machine.

#### RNeasy clean-up of IVT product

10 [276] Follow previous instructions for RNeasy columns or refer to Qiagen's RNeasy protocol handbook.

[277] cRNA will most likely need to be ethanol precipitated. Resuspend in a volume compatible with the fragmentation step.

#### Fragmentation

15 [278] 15 ug of labeled RNA is usually fragmented. Try to minimize the fragmentation reaction volume; a 10 ul volume is recommended but 20 ul is all right. Do not go higher than 20 ul because the magnesium in the fragmentation buffer contributes to precipitation in the hybridization buffer.

20 [279] Fragment RNA by incubation at 94 C for 35 minutes in 1 x Fragmentation buffer.

5 x Fragmentation buffer:

200 mM Tris-acetate, pH 8.1

500 mM KOAc

25 150 mM MgOAc

[280] The labeled RNA transcript can be analyzed before and after fragmentation. Samples can be heated to 65C for 15 minutes and electrophoresed on 1% agarose/TBE gels to get an approximate idea of the transcript size range

30

#### Hybridization

[281] 200 ul (10ug cRNA) of a hybridization mix is put on the chip. If multiple hybridizations are to be done (such as cycling through a 5 chip set), then it is recommended that an initial hybridization mix of 300 ul or more be made.

Hybridization Mix: fragment labeled RNA (50ng/ul final conc.)

50 pM 948-b control oligo

1.5 pM BioB

5 5 pM BioC

25 pM BioD

100 pM CRE

0.1mg/ml herring sperm DNA

0.5mg/ml acetylated BSA

10 to 300 ul with 1xMES hyb. buffer

[282] The instruction manuals for the products used herein are incorporated herein in their entirety.

15 Labeling Protocol Provided Herein

Hybridization reaction:

Start with non-biotinylated IVT (purified by RNeasy columns)

(see example 1 for steps from tissue to IVT)

IVT antisense RNA; 4 µg: µl

20 Random Hexamers (1 µg/µl): 4 µl

H<sub>2</sub>O: µl

14 µl

25 - Incubate 70°C, 10 min. Put on ice.

Reverse transcription:

5X First Strand (BRL) buffer: 6 µl

0.1 M DTT: 3 µl

30 50X dNTP mix: 0.6 µl

H<sub>2</sub>O: 2.4 µl

Cy3 or Cy5 dUTP (1mM): 3 µl

SS RT II (BRL): 1 µl

16 µl

- Add to hybridization reaction.
- Incubate 30 min., 42°C.
- Add 1 µl SSII and let go for another hour.

Put on ice.

- 5           - 50X dNTP mix (25mM of cold dATP, dCTP, and dGTP, 10mM of dTTP: 25 µl each of 100mM dATP, dCTP, and dGTP; 10 µl of 100mM dTTP to 15 µl H<sub>2</sub>O. dNTPs from Pharmacia)

**RNA degradation:**

- 10           86 µl H<sub>2</sub>O
- Add 1.5 µl 1M NaOH/ 2mM EDTA, incubate at 65°C, 10 min.
- 10 µl 10N NaOH
- 4 µl 50mM EDTA
- U-Con 30
- 15           500 µl TE/sample spin at 7000g for 10 min, save flow through for purification

**Qiagen purification:**

- suspend u-con recovered material in 500µl buffer PB
  - proceed w/ normal Qiagen protocol
- 20           DNase digest:
- Add 1 µl of 1/100 dil of DNase/30µl Rx and incubate at 37°C for 15 min.
  - 5 min 95°C to denature enzyme

**Sample preparation:**

- 25           - Add:
- Cot-1 DNA: 10 µl
- 50X dNTPs: 1 µl
- Na pyro phosphate: 7.5 µl
- 10mg/ml Herring sperm DNA 1ul of 1/10 dilution
- 30           21.8 final vol.
- Dry down in speed vac.
  - Resuspend in 15 µl H<sub>2</sub>O.
  - Add 0.38 µl 10% SDS.
  - Heat 95°C, 2 min.

- Slow cool at room temp. for 20 min.

Put on slide and hybridize overnight at 64°C.

5                    **Washing after the hybridization:**

3X SSC/0.03% SDS:                    2 min. 37.5 ml 20X SSC+0.75ml 10% SDS in  
250ml H<sub>2</sub>O

1X SSC: 5 min.                    12.5 ml 20X SSC in 250ml H<sub>2</sub>O

0.2X SSC: 5 min.                    2.5 ml 20X SSC in 250ml H<sub>2</sub>O

10                    Dry slides in centrifuge, 1000 RPM, 1min.

[283] Scan using appropriate Photomultiplier tube (PMT) and fluorescent  
excitation and emission channels.

[284] The results are shown in Table 1 and Table 2. The lists of genes come  
from colorectal tumors from a variety of stages of the disease. The genes that are up  
15 regulated in the tumors (overall) were also found to be expressed at a limited amount or not at  
all in the body map. The body map consists of at least 28 tissue types, including Adrenal  
Gland, Bladder, Bone Marrow, Brain, Breast, Cervix, Colon, Diaphragm, Heart, Kidney,  
Liver, Lung, Lymph Node, Muscle, Pancreas, Prostate, Rectum, Salivary Gland, Skin, Small  
Intestine, Spinal Cord, Spleen, Stomach, Testis, Thymus, Thyroid Trachea and Uterus. As  
20 indicated, some of the Accession numbers include expression sequence tags (ESTs). Thus, in  
one embodiment herein, genes within an expression profile, also termed expression profile  
genes, include ESTs and are not necessarily full length.

[285] Table 1 shows Accession numbers for 1747 genes upregulated in colon  
tumor tissue. The table provides the exemplar accession numbers, Unigene ID numbers,  
25 unique Eos codes, descriptions of the genes encoded, and relative amount of expression as  
compared with expression in other normal body tissue.

**TABLE 1. GENES INVOLVED IN COLORECTAL CANCER**

| 5  | PKey      | Primekey(unique probeset identifier) |          |           |   |                   |
|----|-----------|--------------------------------------|----------|-----------|---|-------------------|
|    | Ex. Accn. | Exemplar accession number            |          |           |   |                   |
|    | Probeset  | Eos Code number                      |          |           |   |                   |
|    | Unigene#  | Unigene number                       |          |           |   |                   |
|    |           |                                      |          |           |   |                   |
| 10 | Pkey      | Probeset                             | Ex Accn  | UniG ID   | UniGene Title   | Ratio TumMet/Body |
| 15 | 332264    | EOS32195                             | N72849   | Hs.115263 | epiregulin  | 17.6              |
|    | 332716    | EOS32647                             | L00058   | Hs.79070  | v-myc avian myelocytomatosis viral oncogene homolog   | 15.0              |
|    | 312845    | EOS12776                             | AI911215 | Hs.186555 | ESTs  | 14.3              |
|    | 310257    | EOS10188                             | AW389247 | Hs.148826 | ESTs  | 11.6              |
|    | 322567    | EOS22498                             | AF155108 |           | EST cluster (not in UniGene)  | 11.5              |
| 20 | 331060    | EOS30991                             | N75081   | Hs.21648  | ESTs  | 10.3              |
|    | 322303    | EOS22234                             | W07459   |           | EST cluster (not in UniGene)  | 9.6               |
|    | 301891    | EOS01822                             | AF131855 | Hs.106127 | Homo sapiens clone 25056 mRNA sequence  | 9.5               |
|    | 318524    | EOS18455                             | AW291511 | Hs.253687 | ESTs  | 8.9               |
|    | 314001    | EOS13932                             | AW168495 | Hs.8750   | ESTs  | 7.8               |
| 25 | 331183    | EOS31114                             | T40769   | Hs.8469   | EST   | 7.3               |
|    | 315429    | EOS15360                             | AW009951 | Hs.206892 | ESTs  | 7.3               |
|    | 303344    | EOS03275                             | AA255977 | Hs.250646 | ESTs; Highly similar to ubiquitin-conjugating enzyme [M.musculus]                               | 6.7               |
|    | 313625    | EOS13556                             | AW468402 | Hs.254020 | ESTs  | 6.7               |
|    | 307084    | EOS07015                             | AI160527 |           | EST singleton (not in UniGene) with exon hit  | 6.1               |
| 30 | 314943    | EOS14874                             | AI476797 | Hs.184572 | cell division cycle 2; G1 to S and G2 to M  | 6.1               |
|    | 303753    | EOS03684                             | AW503733 | Hs.170315 | ESTs  | 5.7               |
|    | 315593    | EOS15524                             | AW198103 | Hs.158154 | ESTs  | 5.3               |
|    | 313604    | EOS13535                             | AI745325 | Hs.182286 | ESTs; Moderately similar to !!!! ALU SUBFAMILY SB2 WARNING ENTRY !!!! [H.sapiens]               | 5.1               |
|    | 312319    | EOS12250                             | AA216698 | Hs.180780 | Homo sapiens agrin precursor mRNA; partial cds  | 5.1               |
| 35 | 312614    | EOS12545                             | AI766732 | Hs.201194 | ESTs  | 4.8               |
|    | 323176    | EOS23107                             | AW071648 | Hs.123199 | ESTs  | 4.8               |
|    | 317916    | EOS17847                             | AI565071 | Hs.159983 | ESTs  | 4.7               |
|    | 301846    | EOS01777                             | R20002   | Hs.6823   | ESTs; Weakly similar to intrinsic factor-B12 receptor precursor [H.sapiens]                     | 4.6               |
|    | 311157    | EOS11088                             | AI990122 | Hs.196988 | ESTs  | 4.6               |
| 40 | 332640    | EOS32571                             | AA417152 | Hs.5101   | protein regulator of cytokinesis 1  | 4.6               |
|    | 311728    | EOS11659                             | AW083000 | Hs.184776 | ribosomal protein L23a  | 4.5               |
|    | 313774    | EOS13705                             | AW136836 | Hs.144583 | ESTs  | 4.5               |
|    | 312339    | EOS12270                             | AA524394 |           | EST cluster (not in UniGene)  | 4.4               |
|    | 315369    | EOS15300                             | AA764918 | Hs.256531 | ESTs  | 4.3               |
| 45 | 303756    | EOS03687                             | AI738488 | Hs.115838 | ESTs  | 4.3               |
|    | 301050    | EOS00981                             | AW136973 | Hs.144475 | ESTs; Weakly similar to mitogen inducible gene mlg-2 [H.sapiens]                                | 4.3               |
|    | 300319    | EOS00250                             | AW157646 | Hs.153506 | ESTs; Weakly similar to microtubule-actin crosslinking factor [M.musculus]                      | 4.3               |
|    | 300664    | EOS00595                             | AI444628 | Hs.256809 | ESTs  | 4.3               |
|    | 302655    | EOS02586                             | AJ227892 |           | EST cluster (not in UniGene) with exon hit  | 4.1               |
| 50 | 315175    | EOS15106                             | AI025842 | Hs.152530 | ESTs  | 4.1               |
|    | 330786    | EOS30717                             | D60374   | Hs.258712 | EST   | 4.1               |
|    | 310875    | EOS10808                             | T47764   | Hs.132917 | ESTs  | 4.1               |
|    | 313425    | EOS13356                             | AA745689 | Hs.186838 | ESTs; Weakly similar to similar to zinc finger 5 protein from Gallus gallus; U51640 [H.sapiens] | 4.0               |
|    | 301804    | EOS01735                             | AA581004 |           | EST cluster (not in UniGene) with exon hit  | 4.0               |
| 55 | 332203    | EOS32134                             | H49388   | Hs.102082 | EST   | 3.9               |
|    | 322968    | EOS22899                             | AI905228 |           | EST cluster (not in UniGene)  | 3.8               |
|    | 321524    | EOS21455                             | N79126   |           | EST cluster (not in UniGene)  | 3.8               |
|    | 302476    | EOS02407                             | AF182294 |           | EST cluster (not in UniGene) with exon hit  | 3.8               |
|    | 303295    | EOS03226                             | AA205625 | Hs.208067 | ESTs  | 3.8               |
| 60 | 310016    | EOS09947                             | AW449612 | Hs.152475 | ESTs  | 3.7               |
|    | 324871    | EOS24802                             | AW297755 | Hs.148832 | ESTs  | 3.7               |
|    | 322887    | EOS22818                             | AI986306 | Hs.233460 | ESTs; Weakly similar to KIAA0969 protein [H.sapiens]  | 3.7               |
|    | 313171    | EOS13102                             | N67879   | Hs.157695 | ESTs  | 3.7               |
|    | 321638    | EOS21569                             | AI356352 | Hs.108932 | ESTs  | 3.7               |
| 65 | 320445    | EOS20376                             | R33916   |           | EST cluster (not in UniGene)  | 3.6               |
|    | 302149    | EOS02080                             | AI383794 | Hs.152337 | protein arginine N-methyltransferase 3(hnRNP methyltransferase S. cerevisiae)-like 3            | 3.6               |
|    | 316905    | EOS16836                             | AW138241 | Hs.210846 | ESTs  | 3.6               |
|    | 313166    | EOS13097                             | AI801098 | Hs.151500 | ESTs  | 3.6               |
|    | 323338    | EOS23269                             | R74219   | Hs.23348  | S-phase kinase-associated protein 2 (p45)   | 3.5               |
| 70 | 311434    | EOS11365                             | AW016607 | Hs.201582 | ESTs  | 3.5               |
|    | 312742    | EOS12673                             | AI650363 | Hs.116462 | ESTs  | 3.4               |
|    | 323587    | EOS23518                             | AI905527 | Hs.141901 | ESTs; Moderately similar to !!!! ALU SUBFAMILY SP WARNING ENTRY !!!! [H.sapiens]                | 3.4               |
|    | 317390    | EOS17321                             | AW136551 | Hs.181245 | ESTs  | 3.4               |
|    | 315282    | EOS15213                             | AI222165 | Hs.144923 | ESTs  | 3.4               |
| 75 | 318565    | EOS18496                             | AI440137 | Hs.164989 | ESTs  | 3.4               |
|    | 307586    | EOS07517                             | AI285499 |           | EST singleton (not in UniGene) with exon hit  | 3.4               |
|    | 321052    | EOS20983                             | AW372884 | Hs.240770 | nuclear cap binding protein subunit 2; 20kD   | 3.3               |
|    | 324338    | EOS24269                             | AI138357 | Hs.247514 | ESTs  | 3.3               |
|    | 307517    | EOS07448                             | AI275055 | Hs.164989 | ESTs  | 3.3               |
| 80 | 314852    | EOS14783                             | AI903735 | Hs.137527 | ESTs; Weakly similar to X-linked retinopathy protein [H.sapiens]                                | 3.3               |
|    | 324657    | EOS24588                             | AW451142 | Hs.255628 | ESTs  | 3.2               |
|    | 314912    | EOS14843                             | AI431345 | Hs.161784 | ESTs  | 3.2               |
|    | 324790    | EOS24721                             | AI334367 | Hs.159337 | ESTs  | 3.2               |
|    | 315498    | EOS15429                             | AA628539 | Hs.116252 | ESTs; Moderately similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]                 | 3.2               |
|    | 312857    | EOS12788                             | AA772279 | Hs.126914 | ESTs  | 3.2               |

|    |        |          |  |  |   |     |
|----|--------|----------|--|--|---|-----|
|    | 300762 | EOS00693 | AI497778   | Hs.168053                                  | ESTs  | 3.2 |
|    | 325587 | EOS25518 | c12_hs gjl6682462[ref] gn 1                      | + 126724 126967 ex 7 7 CDSI 2.44 244 3099  | CH.12_hs gjl6682462   | 3.2 |
| 5  | 320654 | EOS20585 | AW263086   | Hs.118112                                  | ESTs  | 3.2 |
|    | 316715 | EOS16646 | AI440266   | Hs.170673                                  | ESTs  | 3.1 |
|    | 333279 | EOS33210 | CH22_522FG_126_1_LINK_EM:AC005500.GENSCAN.8-1    |  | CH22_FGENES.126_1   | 3.1 |
|    | 309689 | EOS09620 | AW236171   | Hs.181357                                  | laminin receptor 1 (67kD; ribosomal protein SA)   | 3.1 |
| 10 | 323846 | EOS23777 | AA337621   | Hs.137635                                  | ESTs  | 3.1 |
|    | 324678 | EOS24609 | AI990739   | Hs.236511                                  | ESTs; Moderately similar to RNA splicing-related protein [R.norvegicus]                 | 3.1 |
|    | 308362 | EOS08293 | AI613519   |  | EST singleton (not in UniGene) with exon hit  | 3.1 |
|    | 308615 | EOS08546 | AI738593   |  | EST singleton (not in UniGene) with exon hit  | 3.0 |
|    | 315397 | EOS15328 | AA218940   | Hs.137516                                  | ESTs  | 3.0 |
| 15 | 302236 | EOS02167 | AI128606   | Hs.167558                                  | zinc finger protein 161   | 3.0 |
|    | 321693 | EOS21624 | AA700017   | Hs.173737                                  | ras-related C3 botulinum toxin substrate 1 (rho family; small GTP binding protein Rac1) | 3.0 |
|    | 330814 | EOS30745 | AA015730   | Hs.247277                                  | ESTs; Weakly similar to transformation-related protein [H.sapiens]                      | 3.0 |
|    | 302977 | EOS02908 | AW263124   |  | EST cluster (not in UniGene) with exon hit  | 3.0 |
|    | 327516 | EOS27447 | c_2_hs gjl6117815[ref] gn 6                      | + 199078 199216 ex 4 4 CDSI 9.15 139 1551  | CH.02_hs gjl6117815   | 2.9 |
| 20 | 333278 | EOS33209 | CH22_521FG_125_2_LINK_EM:AC005500.GENSCAN.7-2    |  | CH22_FGENES.125_2   | 2.9 |
|    | 302088 | EOS02019 | U77629   | Hs.135639                                  | achaete-scute complex (Drosophila) homolog-like 2                                       | 2.9 |
|    | 322718 | EOS22649 | AF150270   | Hs.233322                                  | ESTs; Weakly similar to cDNA EST EMBL:T01156 comes from this gene [C.elegans]           | 2.9 |
| 25 | 329154 | EOS29085 | c_x_hs gjl5868686[ref] gn 2                      | - 200851 201356 ex 1 3 CDSI 30.28 506 1812 | CH.X_hs gjl5868686  | 2.9 |
|    | 315978 | EOS15909 | AA830893   | Hs.119769                                  | ESTs  | 2.9 |
|    | 302677 | EOS02608 | H63227   | Hs.132880                                  | ESTs; Highly similar to ubiquitin-conjugating enzyme [M.musculus]                       | 2.9 |
|    | 315007 | EOS14938 | AI806583   | Hs.125291                                  | ESTs  | 2.9 |
| 30 | 303780 | EOS03711 | AI424014   | Hs.243450                                  | ESTs; Moderately similar to KIAA0456 protein [H.sapiens]                                | 2.9 |
|    | 331362 | EOS31293 | AA417956   | Hs.40782                                   | ESTs  | 2.9 |
|    | 335815 | EOS35746 | CH22_3187FG_618_3_LINK_EM:AC005500.GENSCAN.510-3 |  | CH22_FGENES.618_3   | 2.8 |
|    | 332070 | EOS32001 | AA598545   | Hs.228138                                  | EST   | 2.8 |
| 35 | 315720 | EOS15651 | AW291875   | Hs.163900                                  | ESTs  | 2.8 |
|    | 311913 | EOS11844 | AI358522   | Hs.221417                                  | ESTs  | 2.8 |
|    | 331014 | EOS30945 | H98597   | Hs.30340                                   | ESTs  | 2.8 |
|    | 322035 | EOS21966 | AL137517   |  | EST cluster (not in UniGene)  | 2.8 |
|    | 338057 | EOS37988 | CH22_6558FG_LINK_EM:AC005500.GENSCAN.160-1       |  | CH22_EM:AC005500.GENSCAN.160-1  | 2.8 |
| 40 | 335829 | EOS35760 | CH22_3202FG_620_3_LINK_EM:AC005500.GENSCAN.512-3 |  | CH22_FGENES.620_3   | 2.8 |
|    | 312136 | EOS12067 | AW451469   | Hs.209990                                  | ESTs  | 2.8 |
|    | 303132 | EOS03063 | AI929819   | Hs.193330                                  | ESTs  | 2.8 |
| 45 | 317548 | EOS17479 | AI654187   | Hs.195704                                  | ESTs  | 2.8 |
|    | 325585 | EOS25516 | c12_hs gjl6682462[ref] gn 1                      | + 73476 73574 ex 5 7 CDSI 8.52 99 309      | CH.12_hs gjl6682462   | 2.7 |
|    | 334631 | EOS34562 | CH22_1939FG_416_7_LINK_EM:AC005500.GENSCAN.277-7 |  | CH22_FGENES.416_7   | 2.7 |
| 50 | 329156 | EOS29087 | c_x_hs gjl5868686[ref] gn 2                      | - 202013 202341 ex 3 3 CDSf 10.23 329 1814 | CH.X_hs gjl5868686  | 2.7 |
|    | 318615 | EOS18546 | AI133617   | Hs.191088                                  | ESTs  | 2.7 |
|    | 300734 | EOS00665 | AW205197   | Hs.240951                                  | ESTs  | 2.7 |
|    | 324430 | EOS24361 | AA464018   |  | EST cluster (not in UniGene)  | 2.7 |
| 55 | 322296 | EOS22227 | W76326   | Hs.251937                                  | ESTs  | 2.7 |
|    | 303842 | EOS03773 | AI337304   | Hs.126268                                  | ESTs; Weakly similar to similar to PDZ domain [C.elegans]                               | 2.7 |
|    | 320909 | EOS20840 | D62269   |  | EST cluster (not in UniGene)  | 2.7 |
|    | 325195 | EOS25126 | T20258   | Hs.171443                                  | ESTs; Weakly similar to actin binding protein MAYVEN [H.sapiens]                        | 2.7 |
|    | 324959 | EOS24890 | AW367745   | Hs.143137                                  | ESTs  | 2.7 |
| 60 | 309997 | EOS09928 | AI291621   | Hs.145199                                  | ESTs  | 2.7 |
|    | 329367 | EOS29298 | c_x_hs gjl5868842[ref] gn 1                      | - 87201 87587 ex 1 4 CDSI 8.13 387 3908    | CH.X_hs gjl5868842  | 2.7 |
|    | 316697 | EOS16628 | AW293174   | Hs.252627                                  | ESTs  | 2.7 |
|    | 313600 | EOS13531 | AA429564   | Hs.185802                                  | ESTs  | 2.7 |
| 65 | 301471 | EOS01402 | AA995014   | Hs.129544                                  | ESTs; Weakly similar to ORF YLL027w [S.cerevisiae]                                      | 2.6 |
|    | 300810 | EOS00741 | AI076890   | Hs.186949                                  | ESTs  | 2.6 |
|    | 319976 | EOS19907 | N48809   | Hs.250824                                  | ESTs  | 2.6 |
|    | 313434 | EOS13365 | W92070   | Hs.231902                                  | ESTs  | 2.6 |
|    | 333849 | EOS33780 | CH22_1118FG_290_8_LINK_EM:AC005500.GENSCAN.146-7 |  | CH22_FGENES.290_8   | 2.6 |
| 70 | 330744 | EOS30675 | AA406142   | Hs.12393                                   | dTDP-D-glucose 4;6-dehydratase  | 2.6 |
|    | 309398 | EOS09329 | AW081820   |  | EST singleton (not in UniGene) with exon hit  | 2.6 |
|    | 338727 | EOS38658 | CH22_7523FG_LINK_EM:AC005500.GENSCAN.500-2       |  | CH22_EM:AC005500.GENSCAN.500-2  | 2.6 |
|    | 324620 | EOS24551 | AA448021   |  | EST cluster (not in UniGene)  | 2.6 |
| 75 | 335755 | EOS35686 | CH22_3122FG_604_4_LINK_EM:AC005500.GENSCAN.493-9 |  | CH22_FGENES.604_4   | 2.6 |
|    | 315858 | EOS15789 | AA737345   |  | EST cluster (not in UniGene)  | 2.6 |
|    | 307288 | EOS07219 | AI205169   |  | EST singleton (not in UniGene) with exon hit  | 2.5 |
| 80 | 330542 | EOS30473 | U23942   | Hs.226213                                  | cytochrome P450; 51 (lanosterol 14-alpha-demethylase)                                   | 2.5 |
|    | 335896 | EOS35827 | CH22_3273FG_635_4_LINK_EM:AC005500.GENSCAN.525-6 |  | CH22_FGENES.635_4   | 2.5 |
|    | 316578 | EOS16509 | AA775623   | Hs.211683                                  | ESTs  | 2.5 |
|    | 329193 | EOS29124 | c_x_hs gjl5868716[ref] gn 3                      | + 168095 168181 ex 9 9 CDSI -1.11 87 2064  | CH.X_hs gjl5868716  | 2.5 |
| 85 | 315193 | EOS15124 | AI241331   | Hs.131765                                  | ESTs  | 2.5 |
|    | 319478 | EOS19409 | R06841   |  | EST cluster (not in UniGene)  | 2.5 |

|    |        |          |   |           |   |
|----|--------|----------|---|-----------|---|
|    | 334727 | EOS34658 | CH22_2038FG_424_1_LINK_EM:AC005500.GENSCAN.285-3                    |           |   |
|    |        |          | CH22_FGENES.424_1   |           | 2.5   |
|    | 328113 | EOS28044 | c_6_hs_gij5868024[ref] gn 2 - 80378 80491 ex 2 3 CDSI 3.89 114 3247 |           |   |
|    |        |          | CH.06_hs_gij5868024   |           | 2.5   |
| 5  | 315214 | EOS15145 | AI915927  | Hs.34771  | ESTs  |
|    | 324718 | EOS24649 | AI557019  | Hs.116467 | ESTs  |
|    | 313326 | EOS13257 | AI088120  | Hs.122329 | ESTs  |
|    | 319480 | EOS19411 | R06933  | Hs.184221 | ESTs  |
|    | 317902 | EOS17833 | AI828602  | Hs.211265 | ESTs  |
| 10 | 323341 | EOS23272 | AI134875  | Hs.192386 | ESTs  |
|    | 336003 | EOS35934 | CH22_3385FG_664_4_LINK_DJ32110.GENSCAN.5-4                          |           |   |
|    |        |          | CH22_FGENES.664_4   |           | 2.5   |
|    | 322992 | EOS22923 | AA142891  | Hs.193165 | ESTs  |
|    | 314911 | EOS14842 | AW292329  | Hs.163481 | ESTs  |
| 15 | 313603 | EOS13534 | AW468119  |           | EST cluster (not in UniGene)  |
|    | 306469 | EOS06400 | AA983792  |           | EST singleton (not in UniGene) with exon hit  |
|    | 324715 | EOS24646 | AI739168  |           | EST cluster (not in UniGene)  |
|    | 302455 | EOS02386 | AA356923  | Hs.240770 | nuclear cap binding protein subunit 2; 20kD   |
|    | 321023 | EOS20954 | H25135  | Hs.125608 | ESTs  |
| 20 | 302099 | EOS02030 | AI021397  | Hs.137576 | ribosomal protein L34 pseudogene 1  |
|    | 314092 | EOS14023 | AI984040  | Hs.226946 | ESTs  |
|    | 318587 | EOS18518 | AA779704  | Hs.168830 | ESTs  |
|    | 303702 | EOS03633 | AW500748  | Hs.224961 | ESTs; Weakly similar to 73 kDa subunit of cleavage and polyadenylation specificity factor [H.sapiens] |
|    | 301822 | EOS01753 | X17033  | Hs.1142   | integrin; alpha 2 (CD49B; alpha 2 subunit of VLA-2 receptor)  |
| 25 | 322694 | EOS22625 | AI110872  |           | EST cluster (not in UniGene)  |
|    | 323333 | EOS23264 | AA228883  |           | EST cluster (not in UniGene)  |
|    | 301954 | EOS01885 | AJ009936  | Hs.118138 | nuclear receptor subfamily 1; group 1; member 2   |
|    | 331363 | EOS31294 | AA421562  | Hs.91011  | anterior gradient 2 (Xenopus laevis) homolog  |
|    | 303811 | EOS03742 | AW182340  | Hs.246155 | ESTs; Weakly similar to DNA TOPOISOMERASE I [H.sapiens]   |
| 30 | 308243 | EOS08174 | AI560037  |           | EST singleton (not in UniGene) with exon hit  |
|    | 336021 | EOS35952 | CH22_3404FG_669_10_LINK_DJ32110.GENSCAN.9-15                        |           |   |
|    |        |          | CH22_FGENES.669_10  |           | 2.4   |
|    | 334789 | EOS34720 | CH22_2101FG_432_14_LINK_EM:AC005500.GENSCAN.293-17                  |           |   |
|    |        |          | CH22_FGENES.432_14  |           | 2.4   |
| 35 | 320807 | EOS20738 | AA086110  | Hs.188536 | Homo sapiens clone 24838 mRNA sequence  |
|    | 328903 | EOS28834 | c_8_hs_gij5868514[ref] gn 1 + 23625 24468 ex 3 5 CDSI 91.18 844 219 |           |   |
|    |        |          | CH.08_hs_gij5868514   |           | 2.4   |
|    | 338759 | EOS38690 | CH22_7581FG_LINK_EM:AC005500.GENSCAN.517-6                          |           |   |
|    |        |          | CH22_EM:AC005500.GENSCAN.517-6                                      |           | 2.3   |
| 40 | 333769 | EOS33700 | CH22_1036FG_271_8_LINK_EM:AC005500.GENSCAN.127-8                    |           |   |
|    |        |          | CH22_FGENES.271_8   |           | 2.3   |
|    | 303597 | EOS03528 | AI792141  | Hs.143560 | ESTs; Weakly similar to brain mitochondrial carrier protein-1 [H.sapiens]                             |
|    | 305898 | EOS05829 | AA872838  | Hs.242463 | keratin 8   |
|    | 304439 | EOS04370 | AA398882  |           | EST singleton (not in UniGene) with exon hit  |
| 45 | 301604 | EOS01535 | AA373124  | Hs.105837 | ESTs; Weakly similar to C17G10.1 [C.elegans]  |
|    | 315071 | EOS15002 | AA552690  | Hs.152423 | ESTs  |
|    | 330565 | EOS30496 | U51095  | Hs.1545   | caudal type homeo box transcription factor 1  |
|    | 331589 | EOS31520 | N71027  | Hs.41856  | ESTs  |
|    | 303216 | EOS03147 | AA581439  | Hs.152328 | ESTs  |
| 50 | 324988 | EOS24919 | T06997  |           | EST cluster (not in UniGene)  |
|    | 312996 | EOS12927 | AA249018  |           | EST cluster (not in UniGene)  |
|    | 332314 | EOS32245 | T25862  | Hs.101774 | ESTs  |
|    | 313325 | EOS13256 | AI420611  | Hs.127832 | ESTs  |
|    | 322991 | EOS22922 | C18965  | Hs.159473 | ESTs  |
| 55 | 335496 | EOS35427 | CH22_2848FG_571_4_LINK_EM:AC005500.GENSCAN.460-25                   |           |   |
|    |        |          | CH22_FGENES.571_4   |           | 2.3   |
|    | 315135 | EOS15066 | AA627561  | Hs.192446 | ESTs  |
|    | 319488 | EOS19419 | AW250340  |           | EST cluster (not in UniGene)  |
|    | 323571 | EOS23502 | AA984133  | Hs.153260 | c-Cbl-interacting protein   |
| 60 | 322826 | EOS22757 | AI807883  | Hs.156932 | ESTs  |
|    | 322221 | EOS22152 | AI890619  | Hs.179662 | nucleosome assembly protein 1-like 1  |
|    | 312242 | EOS12173 | AI380207  | Hs.125276 | ESTs  |
|    | 315238 | EOS15169 | AA593867  | Hs.170890 | ESTs  |
|    | 315168 | EOS15099 | AA622130  | Hs.152524 | ESTs  |
| 65 | 300504 | EOS00435 | AW204624  | Hs.192927 | ESTs; Weakly similar to Lim kinase [H.sapiens]  |
|    | 323243 | EOS23174 | W44372  |           | EST cluster (not in UniGene)  |
|    | 331628 | EOS31559 | R80965  | Hs.204079 | ESTs  |
|    | 320746 | EOS20677 | AA128302  |           | EST cluster (not in UniGene)  |
|    | 324598 | EOS24529 | AA502659  | Hs.163986 | ESTs  |
| 70 | 308667 | EOS08598 | AI758754  |           | EST singleton (not in UniGene) with exon hit  |
|    | 302944 | EOS02875 | AA340708  | Hs.256204 | ESTs; Weakly similar to cyclic nucleotide-gated channel beta subunit [R.norvegicus]                   |
|    | 316291 | EOS16222 | AW375974  | Hs.156704 | ESTs  |
|    | 315296 | EOS15227 | AA876905  | Hs.125286 | ESTs  |
|    | 334150 | EOS34081 | CH22_1429FG_339_1_LINK_EM:AC005500.GENSCAN.189-1                    |           |   |
| 75 |        |          | CH22_FGENES.339_1   |           | 2.2   |
|    | 331380 | EOS31311 | AA453266  | Hs.246131 | ESTs  |
|    | 321795 | EOS21726 | AI796896  | Hs.222446 | ESTs  |
|    | 331493 | EOS31424 | N34357  | Hs.44571  | ESTs  |
|    | 312890 | EOS12821 | AI813654  | Hs.127478 | ESTs  |
| 80 | 315583 | EOS15514 | AW003622  | Hs.126555 | ESTs  |
|    | 314306 | EOS14237 | AI697901  | Hs.192425 | ESTs  |
|    | 314138 | EOS14069 | AA740616  |           | EST cluster (not in UniGene)  |
|    | 302656 | EOS02587 | AW293005  | Hs.220905 | ESTs  |
|    | 313564 | EOS13495 | AA810141  | Hs.192182 | ESTs  |
| 85 | 332792 | EOS32723 | CH22_8FG_3_2_LINK_C4G1.GENSCAN.3-2                                  |           |   |
|    |        |          | CH22_FGENES.3_2   |           | 2.2   |

|    |        |          |  |                     |  |     |
|----|--------|----------|--|---------------------|--|-----|
| 5  | 332020 | EOS31951 | AA488895   | Hs.105219           | ESTs   | 2.2 |
|    | 315143 | EOS15074 | AA878324   | Hs.192734           | ESTs   | 2.2 |
|    | 313385 | EOS13316 | AI032087   | Hs.176711           | ESTs   | 2.2 |
|    | 323835 | EOS23766 | AL042005   |                     | EST cluster (not in UniGene)   | 2.2 |
|    | 314014 | EOS13945 | AW291847   | Hs.121715           | ESTs; Weakly similar to HP protein [H.sapiens]                               | 2.2 |
| 10 | 336016 | EOS35947 | CH22_3399FG_669_5_LINK_DJ32110                                     | GENSCAN.9-10        |  | 2.2 |
|    |        |          |  | CH22_FGENES.669_5   |  | 2.2 |
|    | 323218 | EOS23149 | AF131846   | Hs.13396            | Homo sapiens clone 25028 mRNA sequence                                       | 2.2 |
|    | 338059 | EOS37990 | CH22_6561FG_LINK_EM:AC005500                                       | GENSCAN.160-4       |  | 2.2 |
|    |        |          |  | CH22_EM:AC005500    | GENSCAN.160-4  | 2.2 |
| 15 | 302613 | EOS02544 | AA371059   | Hs.251636           | ubiquitin specific protease 3  | 2.2 |
|    | 304852 | EOS04783 | AA588595   |                     | EST singleton (not in UniGene) with exon hit                                 | 2.2 |
|    | 308457 | EOS08388 | AI669859   |                     | EST singleton (not in UniGene) with exon hit                                 | 2.2 |
|    | 311736 | EOS11667 | AA765897   |                     | EST cluster (not in UniGene)   | 2.2 |
|    | 334183 | EOS34114 | CH22_1464FG_350_13_LINK_EM:AC005500                                | GENSCAN.209-16      |  | 2.2 |
| 20 |        |          |  | CH22_FGENES.350_13  |  | 2.2 |
|    | 315021 | EOS14952 | AA533447   |                     | EST cluster (not in UniGene)   | 2.2 |
|    | 303013 | EOS02944 | F07898   | Hs.214190           | interleukin enhancer binding factor 1  | 2.2 |
|    | 315006 | EOS14937 | AI538613   | Hs.135657           | ESTs   | 2.2 |
|    | 337534 | EOS37465 | CH22_5803FG_828_3_LINK_EM:AC005500                                 | GENSCAN.828-3       |  | 2.2 |
| 25 | 303276 | EOS03207 | AA431599   | Hs.132799           | ESTs   | 2.1 |
|    | 318617 | EOS18548 | AW247252   | Hs.75514            | nucleoside phosphorylase   | 2.1 |
|    | 330760 | EOS30691 | AA448663   | Hs.30469            | ESTs   | 2.1 |
|    | 319545 | EOS19476 | R83716   | Hs.14355            | ESTs   | 2.1 |
|    | 312252 | EOS12183 | AI128388   | Hs.143655           | ESTs   | 2.1 |
| 30 | 322882 | EOS22813 | AW248508   | Hs.2491             | DiGeorge syndrome critical region gene 2                                     | 2.1 |
|    | 312684 | EOS12615 | AW294020   | Hs.117721           | ESTs   | 2.1 |
|    | 315782 | EOS15713 | AW515455   | Hs.115558           | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]  | 2.1 |
|    | 320076 | EOS20007 | AI653733   | Hs.204079           | ESTs   | 2.1 |
|    | 300566 | EOS00497 | H86709   | Hs.21371            | son of sevenless (Drosophila) homolog 1                                      | 2.1 |
| 35 | 300908 | EOS00839 | AA618335   | Hs.146137           | ESTs; Weakly similar to putative [C.elegans]                                 | 2.1 |
|    | 314778 | EOS14709 | AW079559   | Hs.152258           | ESTs   | 2.1 |
|    | 319233 | EOS19164 | R21054   | Hs.211522           | ESTs   | 2.1 |
|    | 335488 | EOS35419 | CH22_2840FG_570_20_LINK_EM:AC005500                                | GENSCAN.460-15      |  | 2.1 |
|    |        |          |  | CH22_FGENES.570_20  |  | 2.1 |
| 40 | 334616 | EOS34547 | CH22_1923FG_411_15_LINK_EM:AC005500                                | GENSCAN.274-22      |  | 2.1 |
|    |        |          |  | CH22_FGENES.411_15  |  | 2.1 |
|    | 306792 | EOS06723 | AI042426   |                     | EST singleton (not in UniGene) with exon hit                                 | 2.1 |
|    | 301661 | EOS01592 | AI815558   |                     | EST cluster (not in UniGene) with exon hit                                   | 2.1 |
|    | 311332 | EOS11263 | AW292247   | Hs.255052           | ESTs   | 2.1 |
| 45 | 314785 | EOS14716 | AI538226   | Hs.135184           | ESTs   | 2.1 |
|    | 301460 | EOS01391 | AW196758   | Hs.165998           | DKFZP564M2423 protein  | 2.1 |
|    | 332015 | EOS31946 | AA487910   | Hs.208800           | ESTs; Weakly similar to !!!! ALU CLASS B WARNING ENTRY !!!! [H.sapiens]      | 2.1 |
|    | 321529 | EOS21460 | AI269506   | Hs.146066           | ESTs   | 2.1 |
|    | 323740 | EOS23671 | AA324643   | Hs.246106           | ESTs   | 2.1 |
| 50 | 336019 | EOS35950 | CH22_3402FG_669_8_LINK_DJ32110                                     | GENSCAN.9-13        |  | 2.1 |
|    |        |          |  | CH22_FGENES.669_8   |  | 2.1 |
|    | 314954 | EOS14885 | AA521381   | Hs.187726           | ESTs   | 2.1 |
|    | 303037 | EOS02968 | AF118395   |                     | EST cluster (not in UniGene) with exon hit                                   | 2.1 |
|    | 302056 | EOS01987 | AI457532   | Hs.126082           | ESTs; Moderately similar to ROSA26AS [M.musculus]                            | 2.1 |
| 55 | 315178 | EOS15109 | AW362945   | Hs.162459           | ESTs   | 2.1 |
|    | 332246 | EOS32177 | N57927   | Hs.120777           | ESTs; Weakly similar to RNA POLYMERASE II ELONGATION FACTOR ELL2 [H.sapiens] | 2.0 |
|    | 334288 | EOS34219 | CH22_1577FG_369_18_LINK_EM:AC005500                                | GENSCAN.229-18      |  | 2.0 |
|    |        |          |  | CH22_FGENES.369_18  |  | 2.0 |
|    | 324690 | EOS24621 | N88286   | Hs.132808           | ESTs; Weakly similar to Similar to S.pombe -rad4+/cut5+product [H.sapiens]   | 2.0 |
| 60 | 305257 | EOS05188 | AA679005   |                     | EST singleton (not in UniGene) with exon hit                                 | 2.0 |
|    | 311315 | EOS11246 | AW450536   | Hs.209260           | ESTs   | 2.0 |
|    | 311988 | EOS11919 | AW016096   | Hs.13801            | ESTs   | 2.0 |
|    | 302638 | EOS02569 | AA463798   | Hs.102696           | ESTs; Weakly similar to C11D2.4 [C.elegans]                                  | 2.0 |
|    | 320531 | EOS20462 | W03691   | Hs.24884            | ESTs; Moderately similar to RNA polymerase I associated factor [M.musculus]  | 2.0 |
| 65 | 323604 | EOS23535 | AI751438   | Hs.182827           | ESTs; Weakly similar to !!!! ALU SUBFAMILY SQ WARNING ENTRY !!!! [H.sapiens] | 2.0 |
|    | 308852 | EOS08783 | AI829848   | Hs.182937           | peptidylprolyl isomerase A (cyclophilin A)                                   | 2.0 |
|    | 320521 | EOS20452 | N31464   | Hs.24743            | ESTs   | 2.0 |
|    | 331306 | EOS31237 | AA252079   | Hs.63931            | dachshund (Drosophila) homolog   | 2.0 |
|    | 314941 | EOS14872 | AA515902   | Hs.130650           | ESTs   | 2.0 |
| 70 | 336684 | EOS36615 | CH22_4167FG_46_1_LINK_EM:AC005500                                  | GENSCAN.46-1        |  | 2.0 |
|    | 301137 | EOS01068 | AF049569   | Hs.137096           | ESTs   | 2.0 |
|    | 338454 | EOS38385 | CH22_7128FG_LINK_EM:AC005500                                       | GENSCAN.360-4       |  | 2.0 |
|    |        |          |  | CH22_EM:AC005500    | GENSCAN.360-4  | 2.0 |
|    | 309700 | EOS09631 | AW241170   | Hs.179661           | Homo sapiens clone 24703 beta-tubulin mRNA; complete cds                     | 2.0 |
| 75 | 330262 | EOS30193 | c_5_p2 gij6671884[gblA gn 1 + 67913 68053 ex 3 3 CDSI 5.41 141 597 |                     |  | 2.0 |
|    |        |          |  | CH.05_p2 gij6671884 |  | 2.0 |
|    | 324163 | EOS24094 | AL046827   | Hs.134651           | ESTs   | 2.0 |
|    | 316493 | EOS16424 | AA766142   | Hs.131810           | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]  | 2.0 |
|    | 311873 | EOS11804 | AA730045   | Hs.187866           | ESTs   | 2.0 |
| 80 | 326757 | EOS26688 | c20_hs gij6249610[refl gn 3 + 74531 74597 ex 1 3 CDSf 9.52 67 1416 |                     |  | 2.0 |
|    |        |          |  | CH.20_hs gij6249610 |  | 2.0 |
|    | 319167 | EOS19098 | F05984   | Hs.250138           | protein phosphatase 2C; magnesium-dependent; catalytic subunit               | 2.0 |
|    | 316011 | EOS15942 | AW516953   | Hs.201372           | ESTs   | 2.0 |
|    | 313635 | EOS13566 | AA507227   | Hs.6390             | ESTs   | 2.0 |
| 85 | 310027 | EOS09958 | AW449009   | Hs.126647           | ESTs   | 2.0 |
|    | 336662 | EOS36593 | CH22_4138FG_41_1_LINK_EM:AC005500                                  | GENSCAN.41-1        |  | 2.0 |
|    | 334648 | EOS34579 | CH22_1956FG_417_15_LINK_EM:AC005500                                | GENSCAN.278-15      |  | 2.0 |
|    |        |          |  | CH22_FGENES.417_15  |  | 2.0 |
|    | 308676 | EOS08607 | AI761036   |                     | EST singleton (not in UniGene) with exon hit                                 | 2.0 |
|    | 312047 | EOS11978 | AA588275   | Hs.14258            | ESTs   | 2.0 |

|    |        |          |                             |  |  |     |
|----|--------|----------|-----------------------------|--|--|-----|
| 5  | 324826 | EOS24757 | AA704806                    | Hs.143842                                  | ESTs   | 2.0 |
|    | 322889 | EOS22820 | AA081924                    | Hs.211417                                  | ESTs   | 2.0 |
|    | 316345 | EOS16276 | AW139408                    | Hs.152940                                  | ESTs   | 2.0 |
|    | 313922 | EOS13853 | AI702038                    | Hs.100057                                  | ESTs   | 2.0 |
|    | 319423 | EOS19354 | T83024                      | Hs.15119                                   | ESTs   | 2.0 |
| 10 | 320244 | EOS20175 | AA296922                    | Hs.129778                                  | gastrointestinal peptide   | 2.0 |
|    | 308957 | EOS08888 | AI869642                    |  | EST singleton (not in UniGene) with exon hit                                     | 2.0 |
|    | 334223 | EOS34154 | CH22_1507FG_360_4_LINK      | EM:AC005500.GENSCAN.218-4                  |  |     |
|    |        |          |                             | CH22_FGENES.360_4                          |  | 1.9 |
|    | 302980 | EOS02911 | W93435                      |  | EST cluster (not in UniGene) with exon hit                                       | 1.9 |
| 15 | 312153 | EOS12084 | AA759250                    | Hs.153028                                  | cytochrome b-561   | 1.9 |
|    | 326460 | EOS26391 | c19_hs gjl5867400[ref] gn 3 | - 142633 142935 ex 1 2 CDSI 19.03 303 1731 |  |     |
|    |        |          |                             | CH.19_hs gjl5867400                        |  | 1.9 |
|    | 319962 | EOS19893 | H06350                      | Hs.135056                                  | ESTs   | 1.9 |
|    | 307064 | EOS06995 | AI149335                    |  | EST singleton (not in UniGene) with exon hit                                     | 1.9 |
| 20 | 331608 | EOS31539 | N89861                      | Hs.44162                                   | ESTs; Weakly similar to cDNA EST yk342h12.5 comes from this gene [C.elegans]     | 1.9 |
|    | 328142 | EOS28073 | c_6_hs gjl5868050[ref] gn 1 | - 9656 9778 ex 2 6 CDSi 11.11 123 3339     |  |     |
|    |        |          |                             | CH.06_hs gjl5868050                        |  | 1.9 |
|    | 312527 | EOS12458 | AI695522                    | Hs.191271                                  | ESTs   | 1.9 |
|    | 318581 | EOS18512 | AA769058                    |  | EST cluster (not in UniGene)   | 1.9 |
| 25 | 319979 | EOS19910 | AB018281                    | Hs.107479                                  | KIAA0738 gene product  | 1.9 |
|    | 336107 | EOS36038 | CH22_3496FG_696_3_LINK      | EM:AC005500.GENSCAN.4-3                    |  |     |
|    |        |          |                             | CH22_FGENES.696_3                          |  | 1.9 |
|    | 305232 | EOS05163 | AA670052                    | Hs.195188                                  | glyceraldehyde-3-phosphate dehydrogenase   | 1.9 |
|    | 315043 | EOS14974 | AA806538                    | Hs.130732                                  | ESTs   | 1.9 |
| 30 | 323377 | EOS23308 | AA133260                    | Hs.8454                                    | protein kinase; cAMP-dependent; regulatory; type II; alpha                       | 1.9 |
|    | 338260 | EOS38191 | CH22_6863FG_LINK            | EM:AC005500.GENSCAN.279-10                 |  |     |
|    |        |          |                             | CH22_EM:AC005500.GENSCAN.279-10            |  | 1.9 |
|    | 334891 | EOS34822 | CH22_2208FG_452_5_LINK      | EM:AC005500.GENSCAN.341-8                  |  |     |
|    |        |          |                             | CH22_FGENES.452_5                          |  | 1.9 |
| 35 | 316055 | EOS15986 | AA693880                    |  | EST cluster (not in UniGene)   | 1.9 |
|    | 312414 | EOS12345 | AI915014                    | Hs.164235                                  | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]      | 1.9 |
|    | 300225 | EOS00156 | AI989963                    | Hs.197505                                  | ESTs   | 1.9 |
|    | 332607 | EOS32538 | R41791                      | Hs.36566                                   | LIM domain kinase 1  | 1.9 |
|    | 312405 | EOS12336 | AI523875                    |  | EST cluster (not in UniGene)   | 1.9 |
| 40 | 313605 | EOS13536 | AI761786                    | Hs.204674                                  | ESTs   | 1.9 |
|    | 337755 | EOS37686 | CH22_6105FG_LINK            | EM:AC000097.GENSCAN.109-2                  |  |     |
|    |        |          |                             | CH22_EM:AC000097.GENSCAN.109-2             |  | 1.9 |
|    | 323216 | EOS23147 | AA332145                    |  | EST cluster (not in UniGene)   | 1.9 |
|    | 334872 | EOS34803 | CH22_2186FG_450_2_LINK      | EM:AC005500.GENSCAN.339-2                  |  |     |
| 45 |        |          |                             | CH22_FGENES.450_2                          |  | 1.9 |
|    | 332034 | EOS31965 | AA489847                    | Hs.112019                                  | ESTs; Moderately similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]  | 1.9 |
|    | 332103 | EOS32034 | AA609161                    | Hs.112657                                  | ESTs; Weakly similar to ORF YOR243c [S.cerevisiae]                               | 1.9 |
|    | 318196 | EOS18127 | AI056776                    | Hs.133397                                  | ESTs   | 1.9 |
|    | 329141 | EOS29072 | c_x_hs gjl6017060[ref] gn 1 | + 343924 343997 ex 2 3 CDSI 8.53 74 1715   |  |     |
| 50 |        |          |                             | CH.X_hs gjl6017060                         |  | 1.9 |
|    | 321539 | EOS21470 | N98619                      | Hs.62461                                   | ARP2 (actin-related protein 2; yeast) homolog                                    | 1.9 |
|    | 313881 | EOS13812 | AA535580                    | Hs.16331                                   | ESTs   | 1.9 |
|    | 314046 | EOS13977 | AW021917                    | Hs.181878                                  | ESTs   | 1.9 |
|    | 336045 | EOS35976 | CH22_3430FG_679_7_LINK      | DJ32110.GENSCAN.18-8                       |  |     |
| 55 |        |          |                             | CH22_FGENES.679_7                          |  | 1.9 |
|    | 324799 | EOS24730 | AW272262                    | Hs.250468                                  | ESTs   | 1.9 |
|    | 312656 | EOS12587 | AW152449                    | Hs.226469                                  | ESTs   | 1.9 |
|    | 324662 | EOS24593 | AW504689                    |  | EST cluster (not in UniGene)   | 1.9 |
|    | 323930 | EOS23861 | AA570698                    | Hs.193203                                  | ESTs   | 1.9 |
| 60 | 314465 | EOS14396 | AA602917                    | Hs.156974                                  | ESTs   | 1.9 |
|    | 335897 | EOS35828 | CH22_3274FG_635_5_LINK      | EM:AC005500.GENSCAN.525-7                  |  |     |
|    |        |          |                             | CH22_FGENES.635_5                          |  | 1.9 |
|    | 321746 | EOS21677 | AI806500                    | Hs.102652                                  | ESTs; Weakly similar to KIAA0437 [H.sapiens]                                     | 1.9 |
|    | 335687 | EOS35618 | CH22_3048FG_596_2_LINK      | EM:AC005500.GENSCAN.488-2                  |  |     |
| 65 |        |          |                             | CH22_FGENES.596_2                          |  | 1.9 |
|    | 330731 | EOS30662 | AA278816                    | Hs.177204                                  | ESTs   | 1.9 |
|    | 315542 | EOS15473 | AA079476                    | Hs.109857                                  | ESTs; Highly similar to CGI-89 protein [H.sapiens]                               | 1.9 |
|    | 336379 | EOS36310 | CH22_3791FG_821_7_LINK      | BA232E17.GENSCAN.4-19                      |  |     |
|    |        |          |                             | CH22_FGENES.821_7                          |  | 1.9 |
| 70 | 305691 | EOS05622 | AA813590                    | Hs.119500                                  | karyopherin alpha 4 (importin alpha 3)   | 1.9 |
|    | 310639 | EOS10570 | AW269082                    | Hs.175162                                  | ESTs   | 1.9 |
|    | 327481 | EOS27412 | c_2_hs gjl5867783[ref] gn 3 | + 104472 104673 ex 1 4 CDSf 14.33 202 1308 |  |     |
|    |        |          |                             | CH.02_hs gjl5867783                        |  | 1.9 |
|    | 301910 | EOS01841 | T84852                      | Hs.98370                                   | cytochrome P540 family member predicted from ESTs                                | 1.9 |
| 75 | 335478 | EOS35409 | CH22_2830FG_569_1_LINK      | EM:AC005500.GENSCAN.456-1                  |  |     |
|    |        |          |                             | CH22_FGENES.569_1                          |  | 1.9 |
|    | 331135 | EOS31066 | R61398                      | Hs.4197                                    | ESTs   | 1.9 |
|    | 335690 | EOS35621 | CH22_3051FG_596_5_LINK      | EM:AC005500.GENSCAN.488-5                  |  |     |
|    |        |          |                             | CH22_FGENES.596_5                          |  | 1.9 |
| 80 | 308047 | EOS07978 | AI459633                    |  | EST singleton (not in UniGene) with exon hit                                     | 1.9 |
|    | 334500 | EOS34431 | CH22_1800FG_397_16_LINK     | EM:AC005500.GENSCAN.260-18                 |  |     |
|    |        |          |                             | CH22_FGENES.397_16                         |  | 1.9 |
|    | 338250 | EOS38181 | CH22_6848FG_LINK            | EM:AC005500.GENSCAN.269-2                  |  |     |
|    |        |          |                             | CH22_EM:AC005500.GENSCAN.269-2             |  | 1.8 |
| 85 | 320618 | EOS20549 | AI220276                    | Hs.235228                                  | EST  | 1.8 |
|    | 335044 | EOS34975 | CH22_2367FG_480_1_LINK      | EM:AC005500.GENSCAN.374-1                  |  |     |
|    |        |          |                             | CH22_FGENES.480_1                          |  | 1.8 |
|    | 313789 | EOS13720 | AI167810                    | Hs.217743                                  | ESTs   | 1.8 |
|    | 311911 | EOS11842 | AI087123                    | Hs.114434                                  | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]      | 1.8 |
|    | 320180 | EOS20111 | AA846203                    | Hs.193974                                  | ESTs; Weakly similar to alternatively spliced product using exon 13A [H.sapiens] | 1.8 |

|    |        |          |  |           |  |     |
|----|--------|----------|--|-----------|--|-----|
| 5  | 311036 | EOS10967 | AI539227   | Hs.214039 | ESTs   | 1.8 |
|    | 323903 | EOS23834 | AA773580   | Hs.193598 | ESTs   | 1.8 |
|    | 318676 | EOS18607 | T57448   | Hs.15467  | ESTs; Moderately similar to putative phosphoinositide 5-phosphatase type II [M.musculus] | 1.8 |
|    | 303007 | EOS02938 | AA478876   | Hs.7037   | pallid (mouse) homolog; pallidin   | 1.8 |
|    | 334806 | EOS34737 | CH22_2119FG_435_7_LINK_EM:AC005500.GENSCAN.296-6                       |           | CH22_FGENES.435_7  | 1.8 |
| 10 | 311767 | EOS11698 | AI076686   | Hs.190066 | ESTs   | 1.8 |
|    | 331750 | EOS31681 | AA284372   | Hs.111471 | ESTs   | 1.8 |
|    | 314872 | EOS14803 | AI144254   | Hs.239726 | ESTs   | 1.8 |
|    | 314071 | EOS14002 | AA192455   | Hs.188690 | ESTs   | 1.8 |
|    | 328450 | EOS28381 | c_7_hs_gij5868425[ref] gn 2 - 209192 209321 ex 2 3 CDSI 10.41 130 1407 |           | CH.07_hs_gij5868425  | 1.8 |
| 15 | 328857 | EOS28788 | c_7_hs_gij6381927[ref] gn 3 - 80557 81051 ex 1 1 CDSO 41.51 495 6090   |           | CH.07_hs_gij6381927  | 1.8 |
|    | 313781 | EOS13712 | AA078836   |           | EST cluster (not in UniGene)   | 1.8 |
|    | 336953 | EOS36884 | CH22_4746FG_361_22_  |           | CH22_FGENES.361-22   | 1.8 |
|    | 300233 | EOS00164 | AI380777   | Hs.189402 | ESTs   | 1.8 |
|    | 326862 | EOS26793 | c20_hs_gij6552465[ref] gn 2 + 107702 107782 ex 12 13 CDSi 3.62 81 2149 |           | CH.20_hs_gij6552465  | 1.8 |
| 20 | 312364 | EOS12295 | R40111   | Hs.187618 | ESTs   | 1.8 |
|    | 321541 | EOS21472 | AI220292   | Hs.254467 | ESTs   | 1.8 |
|    | 307432 | EOS07363 | AI244259   | Hs.181165 | eukaryotic translation elongation factor 1 alpha 1                                       | 1.8 |
|    | 320921 | EOS20852 | R94038   | Hs.199538 | inhibin; beta C  | 1.8 |
|    | 333110 | EOS33041 | CH22_338FG_79_16_LINK_EM:AC000097.GENSCAN.59-15                        |           | CH22_FGENES.79_16  | 1.8 |
| 25 | 324914 | EOS24845 | AA847510   | Hs.161292 | ESTs   | 1.8 |
|    | 312681 | EOS12612 | AI028149   | Hs.193124 | pyruvate dehydrogenase kinase; isoenzyme 3   | 1.8 |
|    | 335697 | EOS35628 | CH22_3058FG_596_12_LINK_EM:AC005500.GENSCAN.488-13                     |           | CH22_FGENES.596_12   | 1.8 |
|    | 308462 | EOS08393 | AI671311   |           | EST singleton (not in UniGene) with exon hit   | 1.8 |
|    | 312138 | EOS12069 | T89405   | Hs.218851 | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]              | 1.8 |
| 30 | 309116 | EOS09047 | AI927149   | Hs.29797  | ribosomal protein L10  | 1.8 |
|    | 320730 | EOS20661 | AA534539   | Hs.151072 | ESTs   | 1.8 |
|    | 300844 | EOS00775 | AL042759   | Hs.191762 | ESTs   | 1.8 |
|    | 337570 | EOS37501 | CH22_5856FG_LINK_C65E1.GENSCAN.4-2                                     |           | CH22_C65E1.GENSCAN.4-2   | 1.8 |
|    | 332756 | EOS32687 | D63479   | Hs.115907 | diacylglycerol kinase; delta (130kD)   | 1.8 |
| 40 | 332161 | EOS32092 | AA621523   | Hs.165464 | ESTs   | 1.8 |
|    | 300942 | EOS00873 | AW275006   | Hs.195969 | ESTs   | 1.8 |
|    | 300680 | EOS00611 | AW468066   | Hs.257712 | ESTs; Weakly similar to KIAA0986 protein [H.sapiens]                                     | 1.8 |
|    | 328783 | EOS28714 | c_7_hs_gij5868309[ref] gn 5 - 73658 73822 ex 2 5 CDSI 0.78 165 5371    |           | CH.07_hs_gij5868309  | 1.8 |
|    | 307542 | EOS07473 | AI280859   |           | EST singleton (not in UniGene) with exon hit   | 1.8 |
| 45 | 331975 | EOS31906 | AA464972   | Hs.99624  | ESTs   | 1.8 |
|    | 321532 | EOS21463 | T77886   | Hs.83428  | nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (p105)              | 1.8 |
|    | 318721 | EOS18652 | Z28504   |           | EST cluster (not in UniGene)   | 1.8 |
|    | 302124 | EOS02055 | AB023967   | Hs.145078 | regulator of differentiation (in S. pombe) 1   | 1.8 |
|    | 323541 | EOS23472 | AI185116   | Hs.104613 | ESTs; Weakly similar to Similar to S.cerevisiae hypothetical protein L3111 [H.sapiens]   | 1.8 |
| 50 | 331057 | EOS30988 | N71399   | Hs.28143  | ESTs   | 1.8 |
|    | 316860 | EOS16791 | AW139099   | Hs.127489 | ESTs   | 1.8 |
|    | 330601 | EOS30532 | U90916   | Hs.82845  | Human clone 23815 mRNA sequence  | 1.8 |
|    | 307334 | EOS07265 | AI214811   | Hs.220615 | ESTs; Weakly similar to TFII-I protein [H.sapiens]                                       | 1.8 |
|    | 323195 | EOS23126 | AI064982   | Hs.117950 | multifunctional polypeptide similar to SAICAR synthetase and AIR carboxylase             | 1.8 |
| 55 | 303856 | EOS03787 | AA968589   | Hs.944    | glucose phosphate isomerase  | 1.8 |
|    | 321553 | EOS21484 | H92449   | Hs.116406 | ESTs   | 1.8 |
|    | 332705 | EOS32636 | T59161   | Hs.76293  | thymosin; beta 10  | 1.8 |
|    | 333139 | EOS33070 | CH22_368FG_83_16_LINK_EM:AC000097.GENSCAN.67-19                        |           | CH22_FGENES.83_16  | 1.8 |
|    | 338997 | EOS38928 | CH22_7881FG_LINK_DA59H18.GENSCAN.8-22                                  |           | CH22_DA59H18.GENSCAN.8-22  | 1.8 |
| 60 | 301509 | EOS01440 | AI026435   | Hs.117532 | ESTs   | 1.8 |
|    | 314522 | EOS14453 | AI732331   | Hs.187750 | ESTs; Moderately similar to !!!! ALU CLASS C WARNING ENTRY !!!! [H.sapiens]              | 1.8 |
|    | 303072 | EOS03003 | AF157833   |           | EST cluster (not in UniGene) with exon hit   | 1.8 |
|    | 305271 | EOS05202 | AA678895   |           | EST singleton (not in UniGene) with exon hit   | 1.8 |
|    | 335287 | EOS35218 | CH22_2629FG_526_11_LINK_EM:AC005500.GENSCAN.420-4                      |           | CH22_FGENES.526_11   | 1.8 |
| 70 | 321286 | EOS21217 | AI380940   |           | EST cluster (not in UniGene)   | 1.8 |
|    | 318740 | EOS18671 | NM_002543  |           | EST cluster (not in UniGene)   | 1.8 |
|    | 323465 | EOS23396 | AA287406   |           | EST cluster (not in UniGene)   | 1.8 |
|    | 300611 | EOS00542 | N75450   |           | EST cluster (not in UniGene) with exon hit   | 1.8 |
|    | 306235 | EOS06166 | AA932299   |           | EST singleton (not in UniGene) with exon hit   | 1.8 |
| 75 | 336721 | EOS36652 | CH22_4244FG_83_17_   |           | CH22_FGENES.83-17  | 1.8 |
|    | 311291 | EOS11222 | AA782601   | Hs.122684 | ESTs   | 1.8 |
|    | 310247 | EOS10178 | AI224982   | Hs.211454 | ESTs   | 1.8 |
|    | 316564 | EOS16495 | AI743571   | Hs.168799 | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]              | 1.8 |
|    | 328170 | EOS28101 | c_6_hs_gij5868071[ref] gn 1 + 93170 93295 ex 9 9 CDSI 13.31 126 3591   |           | CH.06_hs_gij5868071  | 1.8 |
| 80 | 300909 | EOS00840 | AW295479   | Hs.154903 | ESTs; Weakly similar to Abl substrate ena [D.melanogaster]                               | 1.8 |
|    | 330869 | EOS30800 | AA115197   | Hs.183702 | ESTs   | 1.8 |
|    | 311048 | EOS10979 | AA506952   | Hs.210508 | ESTs   | 1.8 |
|    | 333764 | EOS33695 | CH22_1031FG_271_3_LINK_EM:AC005500.GENSCAN.127-3                       |           | CH22_FGENES.271_3  | 1.8 |
|    | 338862 | EOS38793 | CH22_7715FG_LINK_DJ32110.GENSCAN.1-6                                   |           | CH22_DJ32110.GENSCAN.1-6   | 1.8 |
| 85 | 331467 | EOS31398 | N22206   | Hs.43112  | ESTs   | 1.8 |
|    | 327742 | EOS27673 | c_5_hs_gij5867944[ref] gn 3 - 143307 143512 ex 1 3 CDSI 11.07 206 172  |           |  | 1.8 |

|    |        |          |  |  |     |
|----|--------|----------|--|--|-----|
|    |        |          |  | CH.05_hs gij5867944  | 1.8 |
|    | 320955 | EOS20886 | AL049415   | Hs.204290 Homo sapiens mRNA; cDNA DKFZp586N2119 (from clone DKFZp586N2119)                   | 1.8 |
|    | 323589 | EOS23520 | AW390054   | Hs.192843 ESTs   | 1.8 |
|    | 319951 | EOS19882 | AA307665   | Hs.14559 ESTs  | 1.8 |
| 5  | 333763 | EOS33694 | CH22_1030FG_271_2_LINK_EM:AC005500.GENSCAN.127-2                       |  |     |
|    |        |          |  | CH22_FGENES.271_2  | 1.7 |
|    | 331046 | EOS30977 | N66563   | Hs.191358 ESTs   | 1.7 |
|    | 320001 | EOS19932 | AA873350   | EST cluster (not in UniGene)   | 1.7 |
| 10 | 316869 | EOS16800 | AI954880   | Hs.134604 ESTs   | 1.7 |
|    | 310774 | EOS10705 | AW134483   | Hs.164371 ESTs   | 1.7 |
|    | 319379 | EOS19310 | T91443   | Hs.193963 ESTs   | 1.7 |
|    | 321549 | EOS21480 | AA470984   | Hs.161947 ESTs   | 1.7 |
|    | 300823 | EOS00754 | AI863068   | Hs.222665 ESTs; Weakly similar to putative zinc finger protein NY-REN-34 antigen [H.sapiens] | 1.7 |
| 15 | 324228 | EOS24159 | AI798146   | Hs.207780 ESTs   | 1.7 |
|    | 313902 | EOS13833 | AI308165   | Hs.156242 ESTs   | 1.7 |
|    | 308928 | EOS08859 | AI863908   | EST singleton (not in UniGene) with exon hit   | 1.7 |
|    | 333770 | EOS33701 | CH22_1037FG_272_1_LINK_EM:AC005500.GENSCAN.127-10                      |  |     |
|    |        |          |  | CH22_FGENES.272_1  | 1.7 |
| 20 | 316934 | EOS16865 | AI571647   | Hs.146170 ESTs   | 1.7 |
|    | 313219 | EOS13150 | N74924   | Hs.182099 ESTs   | 1.7 |
|    | 317360 | EOS17291 | AI125252   | Hs.126419 ESTs   | 1.7 |
|    | 303530 | EOS03461 | AI274851   | Hs.258744 ESTs   | 1.7 |
|    | 334739 | EOS34670 | CH22_2051FG_424_14_LINK_EM:AC005500.GENSCAN.285-16                     |  |     |
|    |        |          |  | CH22_FGENES.424_14   | 1.7 |
| 25 | 337670 | EOS37601 | CH22_5996FG_LINK_EM:AC000097.GENSCAN.57-2                              |  |     |
|    |        |          |  | CH22_EM:AC000097.GENSCAN.57-2  | 1.7 |
|    | 312079 | EOS12010 | T79745   | Hs.189717 ESTs   | 1.7 |
|    | 320211 | EOS20142 | AL039402   | Hs.125783 DEME-6 protein   | 1.7 |
| 30 | 316218 | EOS16149 | AW207642   | Hs.174021 ESTs   | 1.7 |
|    | 335682 | EOS35613 | CH22_3043FG_595_2_LINK_EM:AC005500.GENSCAN.487-11                      |  |     |
|    |        |          |  | CH22_FGENES.595_2  | 1.7 |
|    | 330696 | EOS30627 | AA022632   | Hs.15825 ESTs  | 1.7 |
|    | 314449 | EOS14380 | AL042667   | Hs.225539 ESTs   | 1.7 |
| 35 | 311972 | EOS11903 | N51511   | Hs.188449 ESTs   | 1.7 |
|    | 307691 | EOS07622 | AI318285   | Hs.182371 prothymosin; alpha (gene sequence 28)  | 1.7 |
|    | 338249 | EOS38180 | CH22_6847FG_LINK_EM:AC005500.GENSCAN.269-1                             |  |     |
|    |        |          |  | CH22_EM:AC005500.GENSCAN.269-1   | 1.7 |
|    | 326399 | EOS26330 | c19_hs gij5867353 ref  gn 1 + 6385 6536 ex 6 6 CDSI 10.69 152 684      |  |     |
| 40 |        |          |  | CH.19_hs gij5867353  | 1.7 |
|    | 313290 | EOS13221 | AI753247   | Hs.206454 ESTs   | 1.7 |
|    | 301615 | EOS01546 | W39477   | EST cluster (not in UniGene) with exon hit   | 1.7 |
|    | 307034 | EOS06965 | AI142526   | EST singleton (not in UniGene) with exon hit   | 1.7 |
|    | 313577 | EOS13508 | AA565051   | Hs.155029 ESTs   | 1.7 |
| 45 | 324703 | EOS24634 | AB009282   | Hs.31086 Homo sapiens mRNA for cytochrome b5; partial cds                                    | 1.7 |
|    | 321317 | EOS21248 | AI937060   | Hs.202040 ESTs; Weakly similar to KIAA0938 protein [H.sapiens]                               | 1.7 |
|    | 312278 | EOS12209 | AW205234   | Hs.201587 ESTs   | 1.7 |
|    | 333358 | EOS33289 | CH22_604FG_141_9_LINK_EM:AC005500.GENSCAN.21-9                         |  |     |
|    |        |          |  | CH22_FGENES.141_9  | 1.7 |
| 50 | 322735 | EOS22666 | AA086123   | EST cluster (not in UniGene)   | 1.7 |
|    | 326752 | EOS26683 | c20_hs gij5867615 ref  gn 1 - 1214 1562 ex 2 2 CDSf 33.07 349 1366     |  |     |
|    |        |          |  | CH.20_hs gij5867615  | 1.7 |
|    | 314733 | EOS14664 | AW452355   | Hs.256037 ESTs   | 1.7 |
|    | 312902 | EOS12833 | AW292797   | Hs.130316 ESTs   | 1.7 |
| 55 | 322653 | EOS22584 | AI828854   | Hs.171891 ESTs   | 1.7 |
|    | 336015 | EOS35946 | CH22_3398FG_669_4_LINK_DJ32110.GENSCAN.9-9                             |  |     |
|    |        |          |  | CH22_FGENES.669_4  | 1.7 |
|    | 324500 | EOS24431 | AW269819   | Hs.169905 ESTs   | 1.7 |
| 60 | 310900 | EOS10831 | AI922728   | Hs.165803 ESTs; Weakly similar to !!!! ALU SUBFAMILY SB WARNING ENTRY !!!! [H.sapiens]       | 1.7 |
|    | 337908 | EOS37839 | CH22_6323FG_LINK_EM:AC005500.GENSCAN.57-1                              |  |     |
|    |        |          |  | CH22_EM:AC005500.GENSCAN.57-1  | 1.7 |
|    | 304084 | EOS04015 | T91986   | EST singleton (not in UniGene) with exon hit   | 1.7 |
|    | 332539 | EOS32470 | AA412528   | Hs.20183 ESTs; Weakly similar to cDNA EST EMBL:T01421 comes from this gene [C.elegans]       | 1.7 |
|    | 314332 | EOS14263 | AL037551   | Hs.95612 ESTs  | 1.7 |
| 65 | 321412 | EOS21343 | AW366305   | EST cluster (not in UniGene)   | 1.7 |
|    | 312187 | EOS12118 | AA700439   | Hs.188490 ESTs   | 1.7 |
|    | 314147 | EOS14078 | AI656135   | Hs.129805 ESTs   | 1.7 |
|    | 303131 | EOS03062 | AW081061   | Hs.103180 actin-like 6   | 1.7 |
|    | 331341 | EOS31272 | AA303125   | Hs.119009 ESTs; Weakly similar to !!!! ALU SUBFAMILY SB2 WARNING ENTRY !!!! [H.sapiens]      | 1.7 |
| 70 | 313615 | EOS13546 | AW295194   | Hs.25264 DKFZP434N126 protein  | 1.7 |
|    | 329598 | EOS29529 | c10_p2 gij3962482 gb A gn 4 + 39924 40220 ex 2 3 CDSi 8.71 297 420     |  |     |
|    |        |          |  | CH.10_p2 gij3962482  | 1.7 |
|    | 303579 | EOS03510 | AA381124   | Hs.193353 ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]        | 1.7 |
|    | 331692 | EOS31623 | W93692   | Hs.47343 ESTs  | 1.7 |
| 75 | 323977 | EOS23908 | AW328177   | Hs.234713 ESTs   | 1.7 |
|    | 332930 | EOS32861 | CH22_151FG_38_4_LINK_C20H12.GENSCAN.29-4                               |  |     |
|    |        |          |  | CH22_FGENES.38_4   | 1.7 |
|    | 326596 | EOS26527 | c19_hs gij6138928 ref  gn 4 + 133386 133563 ex 7 9 CDSi -1.32 178 3520 |  |     |
|    |        |          |  | CH.19_hs gij6138928  | 1.7 |
| 80 | 314946 | EOS14877 | AI097229   | Hs.217484 ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]        | 1.7 |
|    | 315357 | EOS15288 | AA608684   | Hs.121705 ESTs; Moderately similar to !!!! ALU CLASS C WARNING ENTRY !!!! [H.sapiens]        | 1.7 |
|    | 324728 | EOS24659 | AA303024   | EST cluster (not in UniGene)   | 1.7 |
|    | 317501 | EOS17432 | AA931245   | Hs.137097 ESTs   | 1.7 |
|    | 332219 | EOS32150 | N22508   | Hs.139315 ESTs   | 1.7 |
| 85 | 335369 | EOS35300 | CH22_2718FG_543_7_LINK_EM:AC005500.GENSCAN.432-9                       |  |     |
|    |        |          |  | CH22_FGENES.543_7  | 1.7 |
|    | 322417 | EOS22348 | W36286   | Hs.171873 ESTs; Weakly similar to PUTATIVE STEROID DEHYDROGENASE KIK-1 [M.musculus]          | 1.7 |

|    |        |          |   |           |   |     |
|----|--------|----------|---|-----------|---|-----|
|    | 316100 | EOS16031 | AW203986  | Hs.213003 | ESTs  | 1.7 |
|    | 314866 | EOS14797 | AW305124  | Hs.191682 | ESTs  | 1.7 |
|    | 300328 | EOS00259 | AW015860  | Hs.224623 | ESTs  | 1.7 |
| 5  | 315676 | EOS15607 | AW002565  | Hs.136590 | ESTs  | 1.7 |
|    | 314183 | EOS14114 | AA748600  |           | EST cluster (not in UniGene)  | 1.7 |
|    | 321354 | EOS21285 | AA078493  |           | EST cluster (not in UniGene)  | 1.7 |
|    | 311904 | EOS11835 | T86907  | Hs.119371 | ESTs  | 1.7 |
|    | 322890 | EOS22821 | AA082030  |           | EST cluster (not in UniGene)  | 1.7 |
| 10 | 302759 | EOS02690 | AI885815  | Hs.184727 | ESTs  | 1.7 |
|    | 324600 | EOS24531 | AA503297  | Hs.117108 | ESTs  | 1.7 |
|    | 314973 | EOS14904 | AW273128  | Hs.254669 | EST   | 1.7 |
|    | 324432 | EOS24363 | AA464510  |           | EST cluster (not in UniGene)  | 1.7 |
|    | 331520 | EOS31451 | N49068  | Hs.93966  | ESTs  | 1.7 |
| 15 | 306380 | EOS08311 | AI623988  |           | EST singleton (not in UniGene) with exon hit  | 1.7 |
|    | 331010 | EOS30941 | H95039  | Hs.32168  | KIAA0442 protein  | 1.7 |
|    | 325363 | EOS25294 | c12_hs gjl5866920[ref] gn 7 + 700446 700516 ex 6 8 CDSi -6.58 71 113  |           | CH.12_hs gjl5866920   | 1.7 |
|    | 310470 | EOS10401 | AI281848  | Hs.165547 | ESTs  | 1.7 |
| 20 | 330711 | EOS30642 | AA164687  | Hs.177576 | mannosyl (alpha-1,3-)-glycoprotein beta-1;4-N-acetylglucosaminyltransferase; isoenzyme A  | 1.7 |
|    | 332074 | EOS32005 | AA599012  | Hs.22826  | ESTs  | 1.7 |
|    | 309732 | EOS09663 | AW262211  | Hs.5662   | guanine nucleotide binding protein (G protein); beta polypeptide 2-like 1   | 1.6 |
|    | 306337 | EOS06268 | AA954221  | Hs.73742  | ribosomal protein; large; P0  | 1.6 |
|    | 335189 | EOS35120 | CH22_2525FG_507_4_LINK_EM:AC005500.GENSCAN.400-4                      |           | CH22_FGENES.507_4   | 1.6 |
| 25 | 316253 | EOS16184 | AI919537  | Hs.118056 | ESTs  | 1.6 |
|    | 332908 | EOS32839 | CH22_129FG_36_12_LINK_C20H12.GENSCAN.28-9                             |           | CH22_FGENES.36_12   | 1.6 |
|    | 310002 | EOS09933 | AI439096  | Hs.25832  | ESTs  | 1.6 |
| 30 | 332258 | EOS32189 | N68670  | Hs.103808 | ESTs; Weakly similar to RanBPM [H.sapiens]  | 1.6 |
|    | 336182 | EOS36113 | CH22_3576FG_715_2_LINK_DA59H18.GENSCAN.19-3                           |           | CH22_FGENES.715_2   | 1.6 |
|    | 328987 | EOS28918 | c_9_hs gjl5868535[ref] gn 1 - 25705 25764 ex 3 10 CDSi 9.90 60 438    |           | CH.09_hs gjl5868535   | 1.6 |
| 35 | 324481 | EOS24412 | AI916284  | Hs.199671 | ESTs  | 1.6 |
|    | 331406 | EOS31337 | AA610064  | Hs.23440  | KIAA1105 protein  | 1.6 |
|    | 332280 | EOS32211 | R38100  | Hs.106294 | ESTs  | 1.6 |
|    | 332173 | EOS32104 | F09281  | Hs.90424  | ESTs  | 1.6 |
|    | 335739 | EOS35670 | CH22_3102FG_601_10_LINK_EM:AC005500.GENSCAN.491-10                    |           | CH22_FGENES.601_10  | 1.6 |
| 40 | 332104 | EOS32035 | AA609177  | Hs.109363 | ESTs  | 1.6 |
|    | 315033 | EOS14964 | AI493046  | Hs.146133 | ESTs  | 1.6 |
|    | 334740 | EOS34671 | CH22_2052FG_424_15_LINK_EM:AC005500.GENSCAN.285-17                    |           | CH22_FGENES.424_15  | 1.6 |
| 45 | 334783 | EOS34714 | CH22_2095FG_432_8_LINK_EM:AC005500.GENSCAN.293-11                     |           | CH22_FGENES.432_8   | 1.6 |
|    | 308010 | EOS07941 | AI439190  | Hs.181165 | eukaryotic translation elongation factor 1 alpha 1  | 1.6 |
|    | 304521 | EOS04452 | AA464716  |           | EST singleton (not in UniGene) with exon hit  | 1.6 |
| 50 | 318719 | EOS18650 | Z25900  | Hs.18724  | Homo sapiens mRNA; cDNA DKFZp564F093 (from clone DKFZp564F093)  | 1.6 |
|    | 321920 | EOS21851 | N63915  |           | EST cluster (not in UniGene)  | 1.6 |
|    | 315019 | EOS14950 | AA532807  | Hs.105822 | ESTs  | 1.6 |
|    | 320793 | EOS20724 | AL049980  | Hs.184216 | DKFZP564C152 protein  | 1.6 |
|    | 305371 | EOS05302 | AA714180  |           | EST singleton (not in UniGene) with exon hit  | 1.6 |
| 55 | 305054 | EOS04985 | AA634127  | Hs.182426 | ribosomal protein S2  | 1.6 |
|    | 314643 | EOS14574 | AI587502  | Hs.192088 | ESTs  | 1.6 |
|    | 308186 | EOS08117 | AI537940  |           | EST singleton (not in UniGene) with exon hit  | 1.6 |
|    | 319371 | EOS19302 | R00321  | Hs.174928 | ESTs  | 1.6 |
|    | 331700 | EOS31631 | Z40011  | Hs.180582 | ESTs  | 1.6 |
|    | 316955 | EOS16886 | AW203959  | Hs.149532 | ESTs  | 1.6 |
| 60 | 314961 | EOS14892 | AW008061  | Hs.231994 | ESTs  | 1.6 |
|    | 336676 | EOS36607 | CH22_4154FG_43_4  |           | CH22_FGENES.43-4  | 1.6 |
|    | 322801 | EOS22732 | AI831910  | Hs.163734 | ESTs  | 1.6 |
|    | 303363 | EOS03294 | AI964095  | Hs.226801 | ESTs; Weakly similar to DIA-156 protein [H.sapiens]   | 1.6 |
|    | 328105 | EOS28036 | c_6_hs gjl5868020[ref] gn 11 - 301705 301784 ex 4 7 CDSi 5.30 80 3147 |           | CH.06_hs gjl5868020   | 1.6 |
| 65 | 325481 | EOS25412 | c12_hs gjl5866957[ref] gn 3 + 47590 47672 ex 4 7 CDSi 2.69 83 1895    |           | CH.12_hs gjl5866957   | 1.6 |
|    | 315361 | EOS15292 | AI335229  | Hs.122031 | ESTs  | 1.6 |
|    | 324902 | EOS24833 | D31323  | Hs.211188 | ESTs  | 1.6 |
| 70 | 336018 | EOS35949 | CH22_3401FG_669_7_LINK_DJ32110.GENSCAN.9-12                           |           | CH22_FGENES.669_7   | 1.6 |
|    | 308747 | EOS08678 | AI804500  | Hs.181165 | eukaryotic translation elongation factor 1 alpha 1  | 1.6 |
|    | 328251 | EOS28182 | c_6_hs gjl6381891[ref] gn 4 + 124444 124557 ex 2 3 CDSi 0.40 114 4554 |           | CH.06_hs gjl6381891   | 1.6 |
| 75 | 303153 | EOS03084 | U09759  | Hs.8325   | mitogen-activated protein kinase 9  | 1.6 |
|    | 327809 | EOS27740 | c_5_hs gjl5867968[ref] gn 3 + 54610 54761 ex 4 4 CDSi 0.78 152 993    |           | CH.05_hs gjl5867968   | 1.6 |
|    | 314107 | EOS14038 | AA806113  | Hs.189025 | ESTs  | 1.6 |
|    | 300304 | EOS00235 | AI637934  | Hs.224978 | ESTs  | 1.6 |
| 80 | 313009 | EOS12940 | W52010  | Hs.191379 | ESTs  | 1.6 |
|    | 331074 | EOS31005 | R08440  |           | yf19f9.s1 Soares fetal liver spleen 1NfLS Homo sapiens cDNA clone IMAGE:127337 3' similar to contains Alu repetitive element; mRNA sequence | 1.6 |
|    | 335773 | EOS35704 | CH22_3142FG_607_9_LINK_EM:AC005500.GENSCAN.496-4                      |           | CH22_FGENES.607_9   | 1.6 |
| 85 | 334991 | EOS34922 | CH22_2312FG_469_11_LINK_EM:AC005500.GENSCAN.365-11                    |           | CH22_FGENES.469_11  | 1.6 |
|    | 322959 | EOS22890 | AI267606  |           | EST cluster (not in UniGene)  | 1.6 |

|    |        |          |  |           |  |     |
|----|--------|----------|--|-----------|--|-----|
| 5  | 323731 | EOS23662 | AA323414   |           | EST cluster (not in UniGene)   | 1.6 |
|    | 331073 | EOS31004 | R07998   | Hs.18628  | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]                      | 1.6 |
|    | 313573 | EOS13504 | AI076259   | Hs.190337 | ESTs   | 1.6 |
|    | 316949 | EOS16880 | AA856749   | Hs.124620 | ESTs   | 1.6 |
|    | 328084 | EOS28015 | c_6_hs_gij6469819[ref]                             | gn 3      | - 155366 155459 ex 1 4 CDSi 1.23 94 2982<br>CH.06_hs_gij6469819                                  | 1.6 |
| 10 | 331526 | EOS31457 | N49967   | Hs.46624  | ESTs   | 1.6 |
|    | 317987 | EOS17918 | AW138174   | Hs.130651 | ESTs   | 1.6 |
|    | 325594 | EOS25525 | c13_hs_gij5866992[ref]                             | gn 4      | - 470474 470566 ex 2 3 CDSi 8.09 93 68<br>CH.13_hs_gij5866992                                    | 1.6 |
|    | 310848 | EOS10779 | AI459554   | Hs.161286 | ESTs   | 1.6 |
|    | 309268 | EOS09199 | AI985821   | Hs.62954  | ferritin; heavy polypeptide 1  | 1.6 |
| 15 | 304518 | EOS04449 | AA461438   |           | EST singleton (not in UniGene) with exon hit   | 1.6 |
|    | 331065 | EOS30996 | N90584   | Hs.9167   | Homo sapiens clone 25085 mRNA sequence   | 1.6 |
|    | 306501 | EOS06432 | AA987294   |           | EST singleton (not in UniGene) with exon hit   | 1.6 |
|    | 323289 | EOS23220 | AL134235   | Hs.222442 | ESTs   | 1.6 |
|    | 334630 | EOS34561 | CH22_1938FG_416_6_LINK_EM:AC005500.GENSCAN.277-6   |           | CH22_FGENES.416_6  | 1.6 |
| 20 | 302025 | EOS01956 | AI091466   | Hs.127241 | DKFZP564F052 protein   | 1.6 |
|    | 328998 | EOS28929 | c_9_hs_gij5868538[ref]                             | gn 1      | + 40996 41104 ex 1 3 CDSi 11.00 109 480<br>CH.09_hs_gij5868538                                   | 1.6 |
|    | 313197 | EOS13128 | AI738851   | Hs.222487 | ESTs   | 1.6 |
|    | 338763 | EOS38694 | CH22_7585FG_LINK_EM:AC005500.GENSCAN.517-16        |           | CH22_EM:AC005500.GENSCAN.517-16  | 1.6 |
|    | 332247 | EOS32178 | N58172   | Hs.109370 | ESTs   | 1.6 |
| 25 | 316724 | EOS16655 | AA810788   | Hs.123337 | ESTs   | 1.6 |
|    | 303306 | EOS03237 | AA215297   |           | EST cluster (not in UniGene) with exon hit   | 1.6 |
|    | 306336 | EOS06287 | AA954198   |           | EST singleton (not in UniGene) with exon hit   | 1.6 |
|    | 308256 | EOS08187 | AI565498   |           | EST singleton (not in UniGene) with exon hit   | 1.6 |
|    | 307056 | EOS06987 | AI148675   |           | EST singleton (not in UniGene) with exon hit   | 1.6 |
| 30 | 321370 | EOS21301 | AJ227900   |           | EST cluster (not in UniGene)   | 1.6 |
|    | 336262 | EOS36193 | CH22_3661FG_754_9_LINK_DA59H18.GENSCAN.57-11       |           | CH22_FGENES.754_9  | 1.6 |
|    | 335497 | EOS35428 | CH22_2849FG_571_5_LINK_EM:AC005500.GENSCAN.460-26  |           | CH22_FGENES.571_5  | 1.6 |
|    | 309582 | EOS09513 | AW169657   |           | EST singleton (not in UniGene) with exon hit   | 1.6 |
|    | 329563 | EOS29494 | c10_p2_gij3962490[gb]A                             | gn 1      | - 410 635 ex 2 2 CDSi 13.80 226 267<br>CH.10_p2_gij3962490                                       | 1.6 |
| 40 | 332504 | EOS32435 | AA053917   | Hs.15106  | chromosome 14 open reading frame 1   | 1.6 |
|    | 308090 | EOS08021 | AI474601   | Hs.2186   | eukaryotic translation elongation factor 1 gamma   | 1.6 |
|    | 331752 | EOS31683 | AA287312   | Hs.191648 | ESTs   | 1.6 |
|    | 330881 | EOS30812 | AA132986   | Hs.69321  | ESTs; Weakly similar to Similiar to mucin and several other Ser-Thr-rich proteins [S.cerevisiae] | 1.6 |
|    | 315647 | EOS15578 | AA648983   | Hs.212911 | ESTs   | 1.6 |
| 45 | 336766 | EOS36697 | CH22_4341FG_143_20_                                |           | CH22_FGENES.143-20   | 1.6 |
|    | 302592 | EOS02523 | AA294921   | Hs.250811 | v-ral simian leukemia viral oncogene homolog B (ras related; GTP binding protein)                | 1.6 |
|    | 315076 | EOS15007 | AI623817   | Hs.168457 | ESTs   | 1.6 |
|    | 337056 | EOS36987 | CH22_4946FG_441_4_                                 |           | CH22_FGENES.441-4  | 1.6 |
|    | 322175 | EOS22106 | AF085975   |           | EST cluster (not in UniGene)   | 1.6 |
| 50 | 336833 | EOS36764 | CH22_4504FG_242_2_                                 |           | CH22_FGENES.242-2  | 1.6 |
|    | 334902 | EOS34833 | CH22_2219FG_452_16_LINK_EM:AC005500.GENSCAN.341-19 |           | CH22_FGENES.452_16   | 1.6 |
|    | 318671 | EOS18602 | AA188823   | Hs.212621 | ESTs   | 1.6 |
|    | 308064 | EOS07995 | AI469273   | Hs.181165 | eukaryotic translation elongation factor 1 alpha 1   | 1.6 |
|    | 320559 | EOS20490 | AB021981   | Hs.159322 | solute carrier family 35 (UDP-N-acetylglucosamine (UDP-GlcNAc) transporter); member 3            | 1.6 |
| 55 | 317881 | EOS17812 | AI827248   | Hs.224398 | ESTs   | 1.6 |
|    | 313078 | EOS13009 | N49730   |           | EST cluster (not in UniGene)   | 1.6 |
|    | 338689 | EOS38620 | CH22_7464FG_LINK_EM:AC005500.GENSCAN.475-3         |           | CH22_EM:AC005500.GENSCAN.475-3   | 1.6 |
|    | 311804 | EOS11735 | AA135159   | Hs.203349 | ESTs   | 1.6 |
|    | 316359 | EOS16290 | AI472213   | Hs.123415 | ESTs   | 1.6 |
| 60 | 330182 | EOS30113 | c_4_p2_gij5123954[emb]                             | gn 4      | + 120156 120245 ex 2 2 CDSi 4.69 90 11<br>CH.04_p2_gij5123954                                    | 1.6 |
|    | 334718 | EOS34649 | CH22_2028FG_421_29_LINK_EM:AC005500.GENSCAN.282-29 |           | CH22_FGENES.421_29   | 1.6 |
|    | 324196 | EOS24127 | AA405524   | Hs.178000 | ESTs   | 1.6 |
|    | 305350 | EOS05281 | AA706676   |           | EST singleton (not in UniGene) with exon hit   | 1.6 |
|    | 331469 | EOS31400 | N22273   | Hs.39140  | ESTs   | 1.6 |
| 70 | 305715 | EOS05646 | AA826884   |           | EST singleton (not in UniGene) with exon hit   | 1.6 |
|    | 314460 | EOS14391 | AI263231   | Hs.145607 | ESTs   | 1.6 |
|    | 317634 | EOS17565 | AA953088   | Hs.127550 | ESTs   | 1.6 |
|    | 335293 | EOS35224 | CH22_2635FG_527_6_LINK_EM:AC005500.GENSCAN.421-9   |           | CH22_FGENES.527_6  | 1.6 |
|    | 305611 | EOS05542 | AA782331   |           | EST singleton (not in UniGene) with exon hit   | 1.6 |
| 75 | 310430 | EOS10361 | AI670843   | Hs.200257 | ESTs   | 1.6 |
|    | 323696 | EOS23627 | AA641201   | Hs.222051 | ESTs   | 1.6 |
|    | 300610 | EOS00541 | N72596   | Hs.99120  | DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide; Y chromosome                                       | 1.6 |
|    | 327364 | EOS27295 | c_1_hs_gij6552412[ref]                             | gn 2      | - 115235 115396 ex 1 9 CDSi 2.77 162 3007<br>CH.01_hs_gij6552412                                 | 1.6 |
|    | 324848 | EOS24779 | AW021857   |           | EST cluster (not in UniGene)   | 1.6 |
| 80 | 321491 | EOS21422 | H70665   | Hs.183960 | ESTs   | 1.6 |
|    | 336367 | EOS36298 | CH22_3779FG_818_11_LINK_BA232E17.GENSCAN.3-17      |           | CH22_FGENES.818_11   | 1.6 |
|    | 331549 | EOS31480 | N56866   | Hs.237507 | EST  | 1.6 |
|    | 328332 | EOS28263 | c_7_hs_gij5868375[ref]                             | gn 6      | + 280154 280289 ex 3 5 CDSi -1.04 136 516<br>CH.07_hs_gij5868375                                 | 1.5 |
|    | 322817 | EOS22748 | C02420   |           | EST cluster (not in UniGene)   | 1.5 |

|    |        |          |   |  |  |     |
|----|--------|----------|---|--|--|-----|
|    | 303983 | EOS03914 | AW514111  | Hs.181165                                | eukaryotic translation elongation factor 1 alpha 1                                   | 1.5 |
|    | 329434 | EOS29365 | c_y_hs gjl5868883[ref] gn 1                       | - 31124 31263 ex 3 20 CDSi 6.38 140 241  |  |     |
|    |        |          |   | CH.Y_hs gjl5868883                       |  | 1.5 |
| 5  | 338196 | EOS38127 | CH22_6763FG_LINK_EM:AC005500.GENSCAN.235-16       |  |  | 1.5 |
|    |        |          | CH22_EM:AC005500.GENSCAN.235-16                   |  |  | 1.5 |
|    | 308488 | EOS08419 | AI682148  | Hs.179661                                | Homo sapiens clone 24703 beta-tubulin mRNA; complete cds                             | 1.5 |
|    | 314883 | EOS14814 | AW178807  | Hs.246182                                | ESTs   | 1.5 |
|    | 307095 | EOS07026 | AI167910  |  | EST singleton (not in UniGene) with exon hit   | 1.5 |
| 10 | 306953 | EOS06884 | AI124971  |  | EST singleton (not in UniGene) with exon hit   | 1.5 |
|    | 331786 | EOS31717 | AA398539  | Hs.97369                                 | EST  | 1.5 |
|    | 303509 | EOS03440 | AW378236  | Hs.256050                                | ESTs   | 1.5 |
|    | 324515 | EOS24446 | AW501686  | Hs.163539                                | ESTs   | 1.5 |
|    | 339323 | EOS39254 | CH22_8284FG_LINK_BA354112.GENSCAN.23-2            |  |  | 1.5 |
|    |        |          | CH22_BA354112.GENSCAN.23-2                        |  |  | 1.5 |
| 15 | 306563 | EOS06494 | AA995296  |  | EST singleton (not in UniGene) with exon hit   | 1.5 |
|    | 316076 | EOS16007 | AW297895  | Hs.116424                                | ESTs   | 1.5 |
|    | 325622 | EOS25553 | c14_hs gjl5867000[ref] gn 2                       | + 69994 70075 ex 6 8 CDSi 9.40 82 194    |  |     |
|    |        |          |   | CH.14_hs gjl5867000                      |  | 1.5 |
| 20 | 309632 | EOS09563 | AW193261  | Hs.156110                                | Immunoglobulin kappa variable 1D-8   | 1.5 |
|    | 314926 | EOS14857 | AI380838  | Hs.124835                                | ESTs   | 1.5 |
|    | 314458 | EOS14389 | AI217440  | Hs.143873                                | ESTs   | 1.5 |
|    | 335219 | EOS35150 | CH22_2558FG_513_2_LINK_EM:AC005500.GENSCAN.406-2  |  |  | 1.5 |
|    |        |          | CH22_FGENES.513_2                                 |  |  | 1.5 |
| 25 | 301079 | EOS01010 | AA305047  | Hs.183654                                | ESTs; Weakly similar to unknown [S.cerevisiae]                                       | 1.5 |
|    | 334122 | EOS34053 | CH22_1400FG_333_3_LINK_EM:AC005500.GENSCAN.185-27 |  |  | 1.5 |
|    |        |          | CH22_FGENES.333_3                                 |  |  | 1.5 |
|    | 308139 | EOS08070 | AI494477  |  | EST singleton (not in UniGene) with exon hit   | 1.5 |
|    | 317412 | EOS17343 | AI301528  | Hs.132604                                | ESTs   | 1.5 |
| 30 | 315073 | EOS15004 | AW452948  | Hs.257631                                | ESTs   | 1.5 |
|    | 313139 | EOS13070 | AA362113  |  | EST cluster (not in UniGene)   | 1.5 |
|    | 307012 | EOS06943 | AI140212  |  | EST singleton (not in UniGene) with exon hit   | 1.5 |
|    | 322895 | EOS22826 | AW470295  | Hs.192152                                | ESTs   | 1.5 |
|    | 303779 | EOS03710 | AA897296  | Hs.221266                                | ESTs   | 1.5 |
| 35 | 312344 | EOS12275 | AI742618  | Hs.181733                                | ESTs; Weakly similar to nitrilase homolog 1 [H.sapiens]                              | 1.5 |
|    | 323632 | EOS23563 | AL039950  |  | EST cluster (not in UniGene)   | 1.5 |
|    | 332336 | EOS32267 | T96130  | Hs.137551                                | ESTs   | 1.5 |
|    | 304547 | EOS04478 | AA486189  |  | EST singleton (not in UniGene) with exon hit   | 1.5 |
|    | 335692 | EOS35623 | CH22_3053FG_596_7_LINK_EM:AC005500.GENSCAN.488-7  |  |  | 1.5 |
|    |        |          | CH22_FGENES.596_7                                 |  |  | 1.5 |
| 40 | 328333 | EOS28264 | c_7_hs gjl5868375[ref] gn 6                       | + 282506 282664 ex 4 5 CDSi 7.71 159 517 |  |     |
|    |        |          |   | CH.07_hs gjl5868375                      |  | 1.5 |
|    | 304143 | EOS04074 | R88737  |  | EST singleton (not in UniGene) with exon hit   | 1.5 |
|    | 329625 | EOS29556 | c11_p2 gjl4567169[gb]A gn 2                       | - 85893 85984 ex 3 5 CDSi 2.24 92 29     |  |     |
|    |        |          |   | CH.11_p2 gjl4567169                      |  | 1.5 |
| 45 | 329960 | EOS29891 | c16_p2 gjl5091594[gb]A gn 1                       | - 1031 1162 ex 1 3 CDSi 10.75 132 415    |  |     |
|    |        |          |   | CH.16_p2 gjl5091594                      |  | 1.5 |
|    | 318975 | EOS18906 | Z44110  |  | EST cluster (not in UniGene)   | 1.5 |
|    | 321875 | EOS21806 | N49122  |  | EST cluster (not in UniGene)   | 1.5 |
| 50 | 320451 | EOS20382 | R26944  | Hs.180777                                | Homo sapiens mRNA; cDNA DKFZp564M0264 (from clone DKFZp564M0264)                     | 1.5 |
|    | 336020 | EOS35951 | CH22_3403FG_669_9_LINK_DJ32110.GENSCAN.9-14       |  |  | 1.5 |
|    |        |          | CH22_FGENES.669_9                                 |  |  | 1.5 |
|    | 332581 | EOS32512 | T28799  | Hs.2913                                  | EphB3  | 1.5 |
|    | 338622 | EOS38553 | CH22_7384FG_LINK_EM:AC005500.GENSCAN.451-1        |  |  | 1.5 |
|    |        |          | CH22_EM:AC005500.GENSCAN.451-1                    |  |  | 1.5 |
| 55 | 330397 | EOS30328 | D14659  | Hs.154387                                | KIAA0103 gene product  | 1.5 |
|    | 314359 | EOS14290 | AA205569  | Hs.194193                                | ESTs   | 1.5 |
|    | 313456 | EOS13387 | AW380579  | Hs.209657                                | ESTs   | 1.5 |
|    | 318486 | EOS18417 | H09123  | Hs.139258                                | ESTs   | 1.5 |
| 60 | 318175 | EOS18106 | AA644624  |  | EST cluster (not in UniGene)   | 1.5 |
|    | 335684 | EOS35615 | CH22_3045FG_595_4_LINK_EM:AC005500.GENSCAN.487-13 |  |  | 1.5 |
|    |        |          | CH22_FGENES.595_4                                 |  |  | 1.5 |
|    | 327814 | EOS27745 | c_5_hs gjl5867968[ref] gn 6                       | + 69377 70566 ex 1 2 CDSi 86.15 1190 999 |  |     |
|    |        |          |   | CH.05_hs gjl5867968                      |  | 1.5 |
| 65 | 322120 | EOS22051 | W84351  | Hs.213846                                | ESTs   | 1.5 |
|    | 311749 | EOS11680 | R06249  | Hs.13911                                 | ESTs   | 1.5 |
|    | 329797 | EOS29728 | c14_p2 gjl6523160[emb] gn 1                       | - 10616 10894 ex 3 6 CDSi 5.86 279 1549  |  |     |
|    |        |          |   | CH.14_p2 gjl6523160                      |  | 1.5 |
| 70 | 330630 | EOS30561 | X78669  | Hs.79088                                 | reticulocalbin 2; EF-hand calcium binding domain                                     | 1.5 |
|    | 303777 | EOS03708 | AA348491  |  | EST cluster (not in UniGene) with exon hit   | 1.5 |
|    | 309656 | EOS09587 | AW197060  | Hs.195188                                | glyceraldehyde-3-phosphate dehydrogenase   | 1.5 |
|    | 326165 | EOS26096 | c17_hs gjl5867208[ref] gn 2                       | - 62787 62929 ex 1 10 CDSi 0.87 143 2037 |  |     |
|    |        |          |   | CH.17_hs gjl5867208                      |  | 1.5 |
| 75 | 308328 | EOS08259 | AI590571  | Hs.186412                                | EST  | 1.5 |
|    | 300601 | EOS00532 | AI762130  | Hs.165619                                | ESTs   | 1.5 |
|    | 303610 | EOS03541 | AA323288  |  | EST cluster (not in UniGene) with exon hit   | 1.5 |
|    | 307856 | EOS07787 | AI366158  |  | EST singleton (not in UniGene) with exon hit   | 1.5 |
|    | 319920 | EOS19851 | R54575  | Hs.13337                                 | ESTs; Weakly similar to similar to Phosphoglucosyltransferase and phosphomannomutase | 1.5 |
|    |        |          |   |  | phosphoserine [C.elegans]  | 1.5 |
| 80 | 332167 | EOS32098 | D57389  | Hs.75447                                 | ralA binding protein 1   | 1.5 |
|    | 316427 | EOS16358 | AI241019  | Hs.145644                                | ESTs   | 1.5 |
|    | 303886 | EOS03817 | AW365963  |  | EST cluster (not in UniGene) with exon hit   | 1.5 |
|    | 314292 | EOS14223 | AA732590  | Hs.134740                                | ESTs   | 1.5 |
|    | 315408 | EOS15339 | AW273261  | Hs.216292                                | ESTs   | 1.5 |
| 85 | 335698 | EOS35629 | CH22_3059FG_597_1_LINK_EM:AC005500.GENSCAN.489-1  |  |  | 1.5 |
|    |        |          | CH22_FGENES.597_1                                 |  |  | 1.5 |
|    | 315084 | EOS15015 | AI821085  | Hs.187796                                | ESTs   | 1.5 |

|    |        |          |  |   |   |     |
|----|--------|----------|--|---|---|-----|
| 5  | 302299 | EOS02230 | R64632   | Hs.182167                                   | hemoglobin; gamma A   | 1.5 |
|    | 306803 | EOS06734 | A055860  | Hs.193717                                   | interleukin 10  | 1.5 |
|    | 315802 | EOS15733 | AA677540   | Hs.117064                                   | ESTs  | 1.5 |
|    | 326257 | EOS26188 | c17_hs gij5867264 refl  gn 6                       | + 222712 222819 ex 2 2 CDSI 4.46 108 3597   |   |     |
| 10 | 319599 | EOS19530 | H56112   |   | CH.17_hs gij5867264   | 1.5 |
|    | 321891 | EOS21822 | AW157424   | Hs.165954                                   | EST cluster (not in UniGene)  | 1.5 |
|    | 335164 | EOS35095 | CH22_2500FG_502_8_LINK_EM:AC005500.GENSCAN.396-23  |   | ESTs  | 1.5 |
|    | 327133 | EOS27064 | c21_hs gij6682522 refl  gn 1                       | + 38069 38938 ex 2 2 CDSI 63.42 870 1583    | CH22_FGENES.502_8   | 1.5 |
| 15 | 317460 | EOS17391 | AA926980   | Hs.131347                                   | CH.21_hs gij6682522   | 1.5 |
|    | 332344 | EOS32275 | W45574   | Hs.252497                                   | ESTs  | 1.5 |
|    | 328801 | EOS28732 | c_7_hs gij5868321 refl  gn 1                       | - 44492 44609 ex 2 3 CDSI 1.71 118 5525     | ESTs  | 1.5 |
|    | 321677 | EOS21608 | N44545   | Hs.251865                                   | CH.07_hs gij5868321   | 1.5 |
| 20 | 331858 | EOS31789 | AA421163   | Hs.163848                                   | ESTs  | 1.5 |
|    | 309243 | EOS09174 | AI972052   |   | EST singleton (not in UniGene) with exon hit  | 1.5 |
|    | 326213 | EOS26144 | c17_hs gij5867224 refl  gn 3                       | - 60751 60927 ex 1 4 CDSI 2.06 177 2687     | ESTs; Highly similar to phosphoserine aminotransferase [H.sapiens]                                | 1.5 |
|    | 321632 | EOS21563 | AA419617   |   | CH.17_hs gij5867224   | 1.5 |
| 25 | 321424 | EOS21355 | AA057301   |   | EST cluster (not in UniGene)  | 1.5 |
|    | 322465 | EOS22396 | AA137152   | Hs.3784                                     | EST cluster (not in UniGene)  | 1.5 |
|    | 333391 | EOS33322 | CH22_637FG_144_6_LINK_EM:AC005500.GENSCAN.25-6     |   | ESTs  | 1.5 |
|    | 333384 | EOS33315 | CH22_630FG_143_23_LINK_EM:AC005500.GENSCAN.24-17   |   | CH22_FGENES.144_6   | 1.5 |
| 30 | 334784 | EOS34715 | CH22_2096FG_432_9_LINK_EM:AC005500.GENSCAN.293-12  |   | CH22_FGENES.143_23  | 1.5 |
|    | 334078 | EOS34009 | CH22_1356FG_327_33_LINK_EM:AC005500.GENSCAN.181-35 |   | CH22_FGENES.432_9   | 1.5 |
|    | 335158 | EOS35089 | CH22_2494FG_502_2_LINK_EM:AC005500.GENSCAN.396-17  |   | CH22_FGENES.327_33  | 1.5 |
|    | 335062 | EOS34993 | CH22_2388FG_482_17_LINK_EM:AC005500.GENSCAN.376-16 |   | CH22_FGENES.502_2   | 1.5 |
| 35 | 333243 | EOS33174 | CH22_482FG_111_7_LINK_EM:AC000097.GENSCAN.120-6    |   | CH22_FGENES.482_17  | 1.5 |
|    | 306380 | EOS06311 | AA968861   |   | CH22_FGENES.111_7   | 1.5 |
|    | 320809 | EOS20740 | AI540299   |   | EST singleton (not in UniGene) with exon hit  | 1.5 |
|    | 332813 | EOS32744 | CH22_29FG_8_1_LINK_C65E1.GENSCAN.2-2               |   | EST cluster (not in UniGene)  | 1.5 |
| 40 | 335817 | EOS35748 | CH22_3189FG_618_5_LINK_EM:AC005500.GENSCAN.510-5   |   | CH22_FGENES.8_1   | 1.5 |
|    | 319551 | EOS19482 | AA761668   |   | CH22_FGENES.618_5   | 1.5 |
|    | 334472 | EOS34403 | CH22_1771FG_394_3_LINK_EM:AC005500.GENSCAN.257-3   |   | EST cluster (not in UniGene)  | 1.5 |
|    | 333029 | EOS32960 | CH22_255FG_68_3_LINK_EM:AC000097.GENSCAN.40-3      |   | CH22_FGENES.394_3   | 1.5 |
| 50 | 308055 | EOS07986 | AI468091   | Hs.119252                                   | CH22_FGENES.68_3  | 1.5 |
|    | 302882 | EOS02813 | AW403330   |   | tumor protein; translationally-controlled 1   | 1.5 |
|    | 314033 | EOS13964 | AA167125   |   | EST cluster (not in UniGene) with exon hit  | 1.5 |
|    | 324928 | EOS24859 | AI932285   | Hs.160569                                   | EST cluster (not in UniGene)  | 1.5 |
| 55 | 329524 | EOS29455 | c10_p2 gij3983507 gb A gn 6                        | - 38025 38143 ex 3 3 CDSI 2.40 119 170      | ESTs  | 1.5 |
|    | 333131 | EOS33062 | CH22_360FG_83_6_LINK_EM:AC000097.GENSCAN.67-10     |   | CH.10_p2 gij3983507   | 1.5 |
|    | 332085 | EOS32016 | AA600353   | Hs.173933                                   | CH22_FGENES.83_6  | 1.5 |
|    | 305369 | EOS05300 | AA714040   |   | ESTs; Weakly similar to NUCLEAR FACTOR 1/X [H.sapiens]  | 1.5 |
| 60 | 300344 | EOS00275 | AW291487   | Hs.213659                                   | EST singleton (not in UniGene) with exon hit  | 1.5 |
|    | 325071 | EOS25002 | H09693   |   | ESTs  | 1.5 |
|    | 323693 | EOS23624 | AW297758   | Hs.249721                                   | EST cluster (not in UniGene)  | 1.5 |
|    | 321899 | EOS21830 | N55158   | Hs.135252                                   | ESTs  | 1.5 |
| 65 | 331857 | EOS31788 | AA421160   | Hs.9456                                     | SVI/SNF related; matrix associated; actin dependent regulator of chromatin; subfamily a; member 5 | 1.5 |
|    | 334850 | EOS34781 | CH22_2164FG_439_36_LINK_EM:AC005500.GENSCAN.311-13 |   | ESTs  | 1.5 |
|    | 322610 | EOS22541 | AF180919   |   | CH22_FGENES.439_36  | 1.5 |
|    | 335332 | EOS35263 | CH22_2677FG_535_6_LINK_EM:AC005500.GENSCAN.426-6   |   | EST cluster (not in UniGene)  | 1.5 |
| 70 | 307565 | EOS07496 | AI282468   |   | CH22_FGENES.535_6   | 1.5 |
|    | 314140 | EOS14071 | AI216473   | Hs.154297                                   | EST singleton (not in UniGene) with exon hit  | 1.5 |
|    | 323011 | EOS22942 | AA580288   |   | ESTs  | 1.5 |
|    | 325366 | EOS25297 | c12_hs gij5866920 refl  gn 9                       | - 920962 921713 ex 1 8 CDSI 15.95 752 167   | EST cluster (not in UniGene)  | 1.5 |
| 75 | 322306 | EOS22237 | W75935   | Hs.146083                                   | CH.12_hs gij5866920   | 1.5 |
|    | 311034 | EOS10965 | AI564023   | Hs.171467                                   | ESTs  | 1.5 |
|    | 305081 | EOS05012 | AA641638   |   | ESTs; Highly similar to NKG2-D TYPE II INTEGRAL MEMBRANE PROTEIN [H.sapiens]                      | 1.5 |
|    | 322933 | EOS22864 | AA099759   |   | EST singleton (not in UniGene) with exon hit  | 1.5 |
| 80 | 335221 | EOS35152 | CH22_2560FG_513_4_LINK_EM:AC005500.GENSCAN.406-4   |   | EST cluster (not in UniGene)  | 1.5 |
|    | 304948 | EOS04879 | AA613107   |   | CH22_FGENES.513_4   | 1.5 |
|    | 334900 | EOS34831 | CH22_2217FG_452_14_LINK_EM:AC005500.GENSCAN.341-17 |   | EST singleton (not in UniGene) with exon hit  | 1.5 |
|    | 318404 | EOS18335 | AI654108   | Hs.135125                                   | CH22_FGENES.452_14  | 1.5 |
| 85 | 339358 | EOS39289 | CH22_8328FG_LINK_BA354112.GENSCAN.31-3             |   | ESTs  | 1.5 |
|    | 327074 | EOS27005 | c21_hs gij6531965 refl  gn 58                      | + 4039993 4040096 ex 3 4 CDSI 0.68 104 1284 | CH22_BA354112.GENSCAN.31-3  | 1.5 |

|    |        |          |   |   |     |
|----|--------|----------|---|---|-----|
|    |        |          |   | CH.21_hs gij6531965   | 1.5 |
|    | 326054 | EOS25985 | c17_hs gij5867184 ref  gn 2 - 146342 146469 ex 3 4 CDSi 10.00 128 426   |   |     |
|    |        |          |   | CH.17_hs gij5867184   | 1.5 |
| 5  | 326892 | EOS26823 | c20_hs gij6682511 ref  gn 5 + 119424 119500 ex 29 30 CDSi 18.89 77 2313 |   |     |
|    |        |          |   | CH.20_hs gij6682511   | 1.5 |
|    | 328767 | EOS28698 | c_7_hs gij6017031 ref  gn 1 - 35625 35723 ex 4 4 CDSf 5.63 99 5262      |   |     |
|    |        |          |   | CH.07_hs gij6017031   | 1.5 |
|    | 337772 | EOS37703 | CH22_6125FG_LINK_EM:AC000097.GENSCAN.119-11                             |   |     |
| 10 |        |          |   | CH22_EM:AC000097.GENSCAN.119-11   | 1.5 |
|    | 312199 | EOS12130 | AW438602 Hs.191179  | ESTs  | 1.5 |
|    | 303506 | EOS03437 | AA340605 Hs.105887  | ESTs  | 1.5 |
|    | 325176 | EOS25107 | T52843  | EST cluster (not in UniGene)  | 1.5 |
|    | 302023 | EOS01954 | AF060567 Hs.126782  | sushi-repeat protein  | 1.5 |
|    | 305833 | EOS05764 | AA857836 Hs.181165  | eukaryotic translation elongation factor 1 alpha 1                              | 1.5 |
| 15 | 309131 | EOS09062 | AI929175 Hs.119122  | ribosomal protein L13a  | 1.5 |
|    | 334184 | EOS34115 | CH22_1465FG_350_15_LINK_EM:AC005500.GENSCAN.209-17                      |   |     |
|    |        |          |   | CH22_FGENES.350_15  | 1.5 |
|    | 335188 | EOS35119 | CH22_2524FG_507_3_LINK_EM:AC005500.GENSCAN.400-3                        |   |     |
| 20 |        |          |   | CH22_FGENES.507_3   | 1.5 |
|    | 304813 | EOS04744 | AA584540  | EST singleton (not in UniGene) with exon hit                                    | 1.5 |
|    | 315359 | EOS15290 | AA608808 Hs.225118  | ESTs  | 1.5 |
|    | 324434 | EOS24365 | AA707249 Hs.98789   | ESTs  | 1.5 |
|    | 327910 | EOS27841 | c_6_hs gij5868162 ref  gn 1 + 21622 21748 ex 6 7 CDSi 3.69 127 449      |   |     |
|    |        |          |   | CH.06_hs gij5868162   | 1.4 |
| 25 | 335671 | EOS35602 | CH22_3031FG_592_3_LINK_EM:AC005500.GENSCAN.485-4                        |   |     |
|    |        |          |   | CH22_FGENES.592_3   | 1.4 |
|    | 334943 | EOS34874 | CH22_2264FG_465_8_LINK_EM:AC005500.GENSCAN.359-8                        |   |     |
|    |        |          |   | CH22_FGENES.465_8   | 1.4 |
| 30 | 326393 | EOS26324 | c19_hs gij5867341 ref  gn 2 + 41702 41841 ex 5 5 CDSi 20.15 140 504     |   |     |
|    |        |          |   | CH.19_hs gij5867341   | 1.4 |
|    | 305296 | EOS05227 | AA687181  | EST singleton (not in UniGene) with exon hit                                    | 1.4 |
|    | 307243 | EOS07174 | AI199957  | EST singleton (not in UniGene) with exon hit                                    | 1.4 |
|    | 320066 | EOS19997 | AW364885 Hs.112442  | ESTs  | 1.4 |
| 35 | 311465 | EOS11396 | AI758660 Hs.206132  | ESTs  | 1.4 |
|    | 302822 | EOS02753 | AW404176 Hs.111611  | ribosomal protein L27   | 1.4 |
|    | 304987 | EOS04918 | AA618044  | EST singleton (not in UniGene) with exon hit                                    | 1.4 |
|    | 330892 | EOS30823 | AA149579 Hs.118258  | ESTs  | 1.4 |
|    | 333385 | EOS33316 | CH22_631FG_143_24_LINK_EM:AC005500.GENSCAN.24-18                        |   |     |
| 40 |        |          |   | CH22_FGENES.143_24  | 1.4 |
|    | 302626 | EOS02557 | AB021870  | EST cluster (not in UniGene) with exon hit                                      | 1.4 |
|    | 318042 | EOS17973 | AW294522 Hs.149991  | ESTs  | 1.4 |
|    | 339361 | EOS39292 | CH22_8331FG_LINK_BA354112.GENSCAN.32-3                                  |   |     |
|    |        |          |   | CH22_BA354112.GENSCAN.32-3  | 1.4 |
| 45 | 309000 | EOS08931 | AI880489  | EST singleton (not in UniGene) with exon hit                                    | 1.4 |
|    | 306004 | EOS05935 | AA889992  | EST singleton (not in UniGene) with exon hit                                    | 1.4 |
|    | 329539 | EOS29470 | c10_p2 gij3983503 gb U gn 1 - 1 326 ex 1 3 CDSi 41.66 326 212           |   |     |
|    |        |          |   | CH.10_p2 gij3983503   | 1.4 |
|    | 313663 | EOS13594 | AI953261 Hs.169813  | ESTs  | 1.4 |
| 50 | 323538 | EOS23469 | AW247696  | EST cluster (not in UniGene)  | 1.4 |
|    | 337595 | EOS37526 | CH22_5884FG_LINK_C20H12.GENSCAN.8-1                                     |   |     |
|    |        |          |   | CH22_C20H12.GENSCAN.8-1   | 1.4 |
|    | 303149 | EOS03080 | AA312995  | EST cluster (not in UniGene) with exon hit                                      | 1.4 |
|    | 308484 | EOS08415 | AI679292  | EST singleton (not in UniGene) with exon hit                                    | 1.4 |
| 55 | 300912 | EOS00843 | AW138724 Hs.168974  | ESTs  | 1.4 |
|    | 315158 | EOS15089 | AA744438 Hs.142476  | ESTs; Weakly similar to !!!! ALU CLASS D WARNING ENTRY !!!! [H.sapiens]         | 1.4 |
|    | 300462 | EOS00393 | AA746501 Hs.14217   | ESTs  | 1.4 |
|    | 312730 | EOS12661 | AI804372 Hs.208661  | ESTs  | 1.4 |
|    | 316868 | EOS16799 | AI660898 Hs.195602  | ESTs  | 1.4 |
| 60 | 337629 | EOS37560 | CH22_5933FG_LINK_C20H12.GENSCAN.28-35                                   |   |     |
|    |        |          |   | CH22_C20H12.GENSCAN.28-35   | 1.4 |
|    | 332518 | EOS32449 | D16562 Hs.155433  | ATP synthase; H+ transporting; mitochondrial F1 complex; gamma polypeptide 1    | 1.4 |
|    | 337422 | EOS37353 | CH22_5624FG_760_2   | CH22_FGENES.760-2   | 1.4 |
|    | 328835 | EOS28766 | c_7_hs gij5868339 ref  gn 5 + 88053 88461 ex 3 3 CDSi 13.78 409 5775    |   |     |
|    |        |          |   | CH.07_hs gij5868339   | 1.4 |
| 65 | 338282 | EOS38213 | CH22_6897FG_LINK_EM:AC005500.GENSCAN.291-4                              |   |     |
|    |        |          |   | CH22_EM:AC005500.GENSCAN.291-4  | 1.4 |
|    | 337895 | EOS37826 | CH22_6303FG_LINK_EM:AC005500.GENSCAN.56-2                               |   |     |
|    |        |          |   | CH22_EM:AC005500.GENSCAN.56-2   | 1.4 |
| 70 | 320330 | EOS20261 | AF026004 Hs.141660  | chloride channel 2  | 1.4 |
|    | 314302 | EOS14233 | AA813118 Hs.163230  | ESTs  | 1.4 |
|    | 313280 | EOS13211 | AI285537 Hs.222830  | ESTs  | 1.4 |
|    | 333222 | EOS33153 | CH22_459FG_105_2_LINK_EM:AC000097.GENSCAN.109-6                         |   |     |
|    |        |          |   | CH22_FGENES.105_2   | 1.4 |
| 75 | 305726 | EOS05657 | AA828156  | EST singleton (not in UniGene) with exon hit                                    | 1.4 |
|    | 312674 | EOS12605 | AI762475 Hs.151327  | ESTs; Moderately similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens] | 1.4 |
|    | 315869 | EOS15800 | AI033547 Hs.132826  | ESTs  | 1.4 |
|    | 327010 | EOS26941 | c21_hs gij5867664 ref  gn 12 + 941057 941139 ex 9 9 CDSi 7.44 83 790    |   |     |
|    |        |          |   | CH.21_hs gij5867664   | 1.4 |
| 80 | 325892 | EOS25823 | c16_hs gij5867088 ref  gn 1 - 10498 10652 ex 2 3 CDSi 3.94 155 870      |   |     |
|    |        |          |   | CH.16_hs gij5867088   | 1.4 |
|    | 302575 | EOS02506 | AF071164 Hs.249171  | homeo box A11   | 1.4 |
|    | 301970 | EOS01901 | AB028962 Hs.120245  | KIAA1039 protein  | 1.4 |
|    | 332207 | EOS32138 | H61475 Hs.237353  | EST   | 1.4 |
|    | 316024 | EOS15955 | AA707141 Hs.193388  | ESTs  | 1.4 |
| 85 | 314599 | EOS14530 | AW206512 Hs.186996  | ESTs  | 1.4 |
|    | 333585 | EOS33516 | CH22_846FG_203_4_LINK_EM:AC005500.GENSCAN.74-6                          |   |     |

|    |        |          |  |  |     |
|----|--------|----------|--|--|-----|
|    |        |          |  | CH22_FGENES.203_4  | 1.4 |
|    | 324670 | EOS24601 | AI525557   | EST cluster (not in UniGene)   | 1.4 |
|    | 321307 | EOS21238 | R85409   | EST cluster (not in UniGene)   | 1.4 |
| 5  | 335170 | EOS35101 | CH22_2506FG_503_1_LINK_EM:AC005500.GENSCAN.397-1                       | CH22_FGENES.503_1  | 1.4 |
|    | 328274 | EOS28205 | c_7_hs gij5868219[ref] gn 2 - 31244 31439 ex 1 11 CDSi 13.06 196 9     | CH.07_hs gij5868219  | 1.4 |
|    | 336880 | EOS36811 | CH22_4619FG_318_8_   | CH22_FGENES.318-8  | 1.4 |
| 10 | 313825 | EOS13756 | AA215470   | EST cluster (not in UniGene)   | 1.4 |
|    | 318410 | EOS18341 | AI138418 Hs.144935   | ESTs   | 1.4 |
|    | 335361 | EOS35292 | CH22_2710FG_541_11_LINK_EM:AC005500.GENSCAN.431-16                     | CH22_FGENES.541_11   | 1.4 |
|    | 319802 | EOS19733 | AI701489 Hs.202501   | ESTs   | 1.4 |
| 15 | 334769 | EOS34700 | CH22_2081FG_429_4_LINK_EM:AC005500.GENSCAN.290-9                       | CH22_FGENES.429_4  | 1.4 |
|    | 312709 | EOS12640 | AW069181 Hs.141146   | ESTs; Weakly similar to transformation-related protein [H.sapiens]               | 1.4 |
|    | 330004 | EOS29935 | c16_p2 gij6623963[gb]A gn 5 - 78872 78999 ex 2 6 CDSi 19.93 128 728    | CH.16_p2 gij6623963  | 1.4 |
| 20 | 313103 | EOS13034 | AI184303 Hs.143806   | ESTs   | 1.4 |
|    | 326359 | EOS26290 | c18_hs gij5867293[ref] gn 1 + 9436 9494 ex 2 3 CDSi 2.16 59 88         | CH.18_hs gij5867293  | 1.4 |
|    | 305211 | EOS05142 | AA668563   | EST singleton (not in UniGene) with exon hit                                     | 1.4 |
|    | 334628 | EOS34559 | CH22_1936FG_416_4_LINK_EM:AC005500.GENSCAN.277-4                       | CH22_FGENES.416_4  | 1.4 |
| 25 | 326919 | EOS26850 | c21_hs gij6456782[ref] gn 2 - 40486 41046 ex 1 5 CDSi 17.70 561 157    | CH.21_hs gij6456782  | 1.4 |
|    | 315527 | EOS15458 | AI791138 Hs.116768   | ESTs   | 1.4 |
| 30 | 306090 | EOS06021 | AA908609   | EST singleton (not in UniGene) with exon hit                                     | 1.4 |
|    | 303316 | EOS03247 | AF033122 Hs.14125  | p53 regulated PA26 nuclear protein   | 1.4 |
|    | 303642 | EOS03573 | AW299459   | EST cluster (not in UniGene) with exon hit                                       | 1.4 |
|    | 314357 | EOS14288 | AA781795 Hs.122587   | ESTs   | 1.4 |
|    | 337102 | EOS37033 | CH22_5033FG_472_7_   | CH22_FGENES.472-7  | 1.4 |
|    | 304384 | EOS04315 | AA235482 Hs.62954  | ferritin; heavy polypeptide 1  | 1.4 |
| 35 | 315117 | EOS15048 | AA828609 Hs.192044   | ESTs   | 1.4 |
|    | 305750 | EOS05681 | AA835250   | EST singleton (not in UniGene) with exon hit                                     | 1.4 |
|    | 311726 | EOS11657 | AW081766 Hs.253920   | ESTs   | 1.4 |
|    | 326996 | EOS26927 | c21_hs gij5867660[ref] gn 4 - 63212 63404 ex 2 6 CDSi 15.70 193 622    | CH.21_hs gij5867660  | 1.4 |
| 40 | 330257 | EOS30188 | c_5_p2 gij6671881[gb]A gn 2 - 143228 143393 ex 1 9 CDSi 11.31 166 586  | CH.05_p2 gij6671881  | 1.4 |
|    | 323864 | EOS23795 | AA340724 Hs.214028   | ESTs   | 1.4 |
|    | 338204 | EOS38135 | CH22_6773FG__LINK_EM:AC005500.GENSCAN.241-3                            | CH22_EM:AC005500.GENSCAN.241-3   | 1.4 |
| 45 | 314025 | EOS13956 | AI983981 Hs.189114   | ESTs   | 1.4 |
|    | 315974 | EOS15905 | AW029203 Hs.191952   | ESTs   | 1.4 |
|    | 335599 | EOS35530 | CH22_2957FG_581_39_LINK_EM:AC005500.GENSCAN.476-37                     | CH22_FGENES.581_39   | 1.4 |
|    | 335364 | EOS35295 | CH22_2713FG_543_2_LINK_EM:AC005500.GENSCAN.432-4                       | CH22_FGENES.543_2  | 1.4 |
| 50 | 303634 | EOS03565 | AI953377 Hs.169425   | ESTs; Weakly similar to predicted using Genefinder [C.elegans]                   | 1.4 |
|    | 315626 | EOS15557 | AA808598 Hs.35353  | ESTs; Weakly similar to H21P03.2 [C.elegans]                                     | 1.4 |
|    | 329936 | EOS29867 | c16_p2 gij6165200[gb]A gn 4 - 82761 82920 ex 3 4 CDSi 1.15 160 199     | CH.16_p2 gij6165200  | 1.4 |
| 55 | 328632 | EOS28563 | c_7_hs gij5868247[ref] gn 1 + 76734 76853 ex 1 4 CDSf 13.95 120 3764   | CH.07_hs gij5868247  | 1.4 |
|    | 330207 | EOS30138 | c_5_p2 gij6013606[gb]A gn 3 - 109912 110004 ex 2 4 CDSi 6.54 93 174    | CH.05_p2 gij6013606  | 1.4 |
|    | 329919 | EOS29850 | c16_p2 gij6223624[gb]A gn 6 - 103492 103681 ex 1 8 CDSi 6.18 190 93    | CH.16_p2 gij6223624  | 1.4 |
| 60 | 331916 | EOS31847 | AA446131 Hs.124918   | ESTs   | 1.4 |
|    | 317617 | EOS17548 | T58194   | EST cluster (not in UniGene)   | 1.4 |
|    | 331943 | EOS31874 | AA453418 Hs.178272   | ESTs   | 1.4 |
|    | 306413 | EOS06344 | AA973288   | EST singleton (not in UniGene) with exon hit                                     | 1.4 |
| 65 | 313607 | EOS13538 | N94169 Hs.194258   | ESTs; Moderately similar to !!!! ALU SUBFAMILY SC WARNING ENTRY !!!! [H.sapiens] | 1.4 |
|    | 336292 | EOS36223 | CH22_3691FG_783_3_LINK_BA354112.GENSCAN.4-7                            | CH22_FGENES.783_3  | 1.4 |
|    | 330453 | EOS30384 | HG3976-HT4246  | Pou-Domain Dna Binding Factor Pit1, Pituitary-Specific                           | 1.4 |
|    | 324602 | EOS24533 | AA503620 Hs.213239   | ESTs   | 1.4 |
| 70 | 332183 | EOS32114 | H08225 Hs.177181   | ESTs   | 1.4 |
|    | 320032 | EOS19963 | AI699772 Hs.202361   | ESTs; Weakly similar to X-linked retinopathy protein [H.sapiens]                 | 1.4 |
|    | 333156 | EOS33087 | CH22_387FG_89_6_LINK_EM:AC000097.GENSCAN.84-8                          | CH22_FGENES.89_6   | 1.4 |
|    | 334156 | EOS34087 | CH22_1435FG_340_6_LINK_EM:AC005500.GENSCAN.190-7                       | CH22_FGENES.340_6  | 1.4 |
| 75 | 334303 | EOS34234 | CH22_1594FG_373_6_LINK_EM:AC005500.GENSCAN.233-5                       | CH22_FGENES.373_6  | 1.4 |
|    | 325513 | EOS25444 | c12_hs gij6017035[ref] gn 1 - 34295 34490 ex 2 7 CDSi 6.49 196 2471    | CH.12_hs gij6017035  | 1.4 |
| 80 | 302758 | EOS02689 | AA984563   | EST cluster (not in UniGene) with exon hit                                       | 1.4 |
|    | 329557 | EOS29488 | c10_p2 gij3962492[gb]A gn 6 - 53197 53647 ex 2 2 CDSf 37.68 451 247    | CH.10_p2 gij3962492  | 1.4 |
|    | 331717 | EOS31648 | AA190888 Hs.153881   | ESTs; Highly similar to NY-REN-62 antigen [H.sapiens]                            | 1.4 |
|    | 325885 | EOS25816 | c16_hs gij5867087[ref] gn 11 + 193212 193377 ex 1 3 CDSf 43.19 166 792 | CH.16_hs gij5867087  | 1.4 |
| 85 | 312180 | EOS12091 | AA805903 Hs.184371   | ESTs   | 1.4 |
|    | 328882 | EOS28813 | c_7_hs gij6552423[ref] gn 2 - 157669 157826 ex 4 6 CDSi 4.91 158 6200  |  | 1.4 |

|    |        |          |   |   |     |
|----|--------|----------|---|---|-----|
|    |        |          |   | CH.07_hs gij6552423   | 1.4 |
|    | 339028 | EOS38959 | CH22_7925FG__LINK_DA59H18.GENSCAN.22-8                                  |   |     |
|    |        |          | CH22_DA59H18.GENSCAN.22-8   |   | 1.4 |
| 5  | 323497 | EOS23428 | AI523613 Hs.221544  | ESTs  | 1.4 |
|    | 316897 | EOS16828 | AA838114  | EST cluster (not in UniGene)  | 1.4 |
|    | 312479 | EOS12410 | AI950844 Hs.128738  | ESTs; Weakly similar to non-lens beta gamma-crystallin like protein [H.sapiens] | 1.4 |
|    | 338535 | EOS38466 | CH22_7251FG__LINK_EM:AC005500.GENSCAN.404-3                             |   |     |
|    |        |          | CH22_EM:AC005500.GENSCAN.404-3  |   | 1.4 |
| 10 | 312754 | EOS12685 | R99834 Hs.250383  | ESTs  | 1.4 |
|    | 327527 | EOS27458 | c_2_hs gij6381882[ref] gn 2 - 98950 99040 ex 4 8 CDSi 5.78 91 1768      |   |     |
|    |        |          | CH.02_hs gij6381882   |   | 1.4 |
|    | 324714 | EOS24645 | AA574312 Hs.245737  | ESTs  | 1.4 |
|    | 302347 | EOS02278 | AF039400 Hs.194659  | chloride channel; calcium activated; family member 1                            | 1.4 |
| 15 | 338008 | EOS37939 | CH22_6490FG__LINK_EM:AC005500.GENSCAN.127-9                             |   |     |
|    |        |          | CH22_EM:AC005500.GENSCAN.127-9  |   | 1.4 |
|    | 315590 | EOS15521 | AA640637 Hs.225817  | ESTs  | 1.4 |
|    | 320825 | EOS20756 | NM_004751   | EST cluster (not in UniGene)  | 1.4 |
|    | 300930 | EOS00861 | AI289481 Hs.136371  | ESTs  | 1.4 |
| 20 | 335225 | EOS35156 | CH22_2564FG_513_10_LINK_EM:AC005500.GENSCAN.406-9                       |   |     |
|    |        |          | CH22_FGENES.513_10  |   | 1.4 |
|    | 337303 | EOS37234 | CH22_5442FG_681_5_  | CH22_FGENES.681-5   | 1.4 |
|    | 317198 | EOS17129 | AI810384 Hs.128025  | ESTs  | 1.4 |
|    | 308991 | EOS08922 | AI879831  | EST singleton (not in UniGene) with exon hit                                    | 1.4 |
| 25 | 325472 | EOS25403 | c12_hs gij6017034[ref] gn 7 - 289581 289657 ex 2 6 CDSi 4.74 77 1786    |   |     |
|    |        |          | CH.12_hs gij6017034   |   | 1.4 |
|    | 301266 | EOS01197 | AA829774  | EST cluster (not in UniGene) with exon hit                                      | 1.4 |
|    | 330901 | EOS30832 | AA157818 Hs.238380  | Human endogenous retroviral protease mRNA; complete cds                         | 1.4 |
|    | 313406 | EOS13337 | AI248314 Hs.132932  | ESTs  | 1.4 |
| 30 | 301454 | EOS01385 | AI751738  | EST cluster (not in UniGene) with exon hit                                      | 1.4 |
|    | 317269 | EOS17200 | AA906411 Hs.127378  | ESTs  | 1.4 |
|    | 338876 | EOS38807 | CH22_7733FG__LINK_DJ32I10.GENSCAN.4-2                                   |   |     |
|    |        |          | CH22_DJ32I10.GENSCAN.4-2  |   | 1.4 |
| 35 | 328481 | EOS28412 | c_7_hs gij5868449[ref] gn 1 - 8987 9180 ex 4 31 CDSi 10.00 194 2103     |   |     |
|    |        |          | CH.07_hs gij5868449   |   | 1.4 |
|    | 314022 | EOS13953 | AW452420 Hs.248678  | ESTs  | 1.4 |
|    | 307640 | EOS07571 | AI301992  | EST singleton (not in UniGene) with exon hit                                    | 1.4 |
|    | 315541 | EOS15472 | AI168233 Hs.123159  | ESTs; Weakly similar to KIAA0668 protein [H.sapiens]                            | 1.4 |
|    | 315489 | EOS15420 | AA628245 Hs.191847  | ESTs  | 1.4 |
| 40 | 327815 | EOS27746 | c_5_hs gij5867968[ref] gn 6 + 70804 71401 ex 2 2 CDSi 27.99 598 1000    |   |     |
|    |        |          | CH.05_hs gij5867968   |   | 1.4 |
|    | 339319 | EOS39250 | CH22_8280FG__LINK_BA354I12.GENSCAN.22-19                                |   |     |
|    |        |          | CH22_BA354I12.GENSCAN.22-19   |   | 1.4 |
|    | 322564 | EOS22495 | W86440 Hs.118344  | ESTs  | 1.4 |
| 45 | 323812 | EOS23743 | AW081373 Hs.199199  | ESTs  | 1.4 |
|    | 303540 | EOS03471 | AA355607 Hs.173590  | ESTs; Weakly similar to MMSET type I [H.sapiens]                                | 1.4 |
|    | 337902 | EOS37833 | CH22_6314FG__LINK_EM:AC005500.GENSCAN.56-13                             |   |     |
|    |        |          | CH22_EM:AC005500.GENSCAN.56-13  |   | 1.4 |
|    | 335289 | EOS35220 | CH22_2631FG_527_2_LINK_EM:AC005500.GENSCAN.421-2                        |   |     |
|    |        |          | CH22_FGENES.527_2   |   | 1.4 |
| 50 | 327919 | EOS27850 | c_6_hs gij5868165[ref] gn 6 + 547701 547800 ex 14 14 CDSi -0.20 100 505 |   |     |
|    |        |          | CH.06_hs gij5868165   |   | 1.4 |
|    | 337674 | EOS37605 | CH22_6005FG__LINK_EM:AC000097.GENSCAN.67-4                              |   |     |
|    |        |          | CH22_EM:AC000097.GENSCAN.67-4   |   | 1.4 |
| 55 | 320087 | EOS20018 | AF032387 Hs.113265  | small nuclear RNA activating complex; polypeptide 4; 190kD                      | 1.4 |
|    | 334939 | EOS34870 | CH22_2259FG_465_3_LINK_EM:AC005500.GENSCAN.359-3                        |   |     |
|    |        |          | CH22_FGENES.465_3   |   | 1.3 |
|    | 303443 | EOS03374 | AA320525  | EST cluster (not in UniGene) with exon hit                                      | 1.3 |
|    | 325929 | EOS25860 | c16_hs gij5867125[ref] gn 2 - 51715 51996 ex 1 1 CDSi 29.05 282 1594    |   |     |
| 60 |        |          | CH.16_hs gij5867125   |   | 1.3 |
|    | 327745 | EOS27676 | c_5_hs gij6531959[ref] gn 1 - 229066 229124 ex 3 6 CDSi 3.01 59 177     |   |     |
|    |        |          | CH.05_hs gij6531959   |   | 1.3 |
|    | 335166 | EOS35097 | CH22_2502FG_502_10_LINK_EM:AC005500.GENSCAN.396-25                      |   |     |
|    |        |          | CH22_FGENES.502_10  |   | 1.3 |
| 65 | 324497 | EOS24428 | AW152624 Hs.136340  | ESTs  | 1.3 |
|    | 338374 | EOS38305 | CH22_7017FG__LINK_EM:AC005500.GENSCAN.327-1                             |   |     |
|    |        |          | CH22_EM:AC005500.GENSCAN.327-1  |   | 1.3 |
|    | 313601 | EOS13532 | R32458 Hs.257711  | ESTs  | 1.3 |
|    | 321415 | EOS21346 | AI377596 Hs.3337  | transmembrane 4 superfamily member 1  | 1.3 |
| 70 | 305309 | EOS05240 | AA699717  | EST singleton (not in UniGene) with exon hit                                    | 1.3 |
|    | 330447 | EOS30378 | HG3546-HT3744   | Pre-Mrna Splicing Factor Sf2p33, Alt. Splice Form 1                             | 1.3 |
|    | 308578 | EOS08509 | AI708573  | EST singleton (not in UniGene) with exon hit                                    | 1.3 |
|    | 315344 | EOS15275 | AW292176 Hs.245834  | ESTs  | 1.3 |
|    | 330503 | EOS30434 | M55024  | Human cell surface glycoprotein P3.58 mRNA, partial cds                         | 1.3 |
| 75 | 308227 | EOS08158 | AI559126 Hs.195188  | glyceraldehyde-3-phosphate dehydrogenase  | 1.3 |
|    | 332222 | EOS32153 | N28271 Hs.176618  | ESTs  | 1.3 |
|    | 323961 | EOS23892 | AL044428 Hs.207345  | ESTs  | 1.3 |
|    | 314530 | EOS14461 | AI052358 Hs.131741  | ESTs  | 1.3 |
|    | 320503 | EOS20434 | NM_005897   | EST cluster (not in UniGene)  | 1.3 |
| 80 | 306820 | EOS06751 | AI074408  | EST singleton (not in UniGene) with exon hit                                    | 1.3 |
|    | 304165 | EOS04096 | H73265  | EST singleton (not in UniGene) with exon hit                                    | 1.3 |
|    | 324302 | EOS24233 | AA543008 Hs.136806  | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]     | 1.3 |
|    | 319128 | EOS19059 | AA393820  | EST cluster (not in UniGene)  | 1.3 |
|    | 317092 | EOS17023 | AI286162 Hs.125657  | ESTs  | 1.3 |
| 85 | 304998 | EOS04929 | AA621203  | EST singleton (not in UniGene) with exon hit                                    | 1.3 |
|    | 331433 | EOS31364 | H68097 Hs.161023  | EST   | 1.3 |
|    | 333348 | EOS33279 | CH22_594FG_140_2_LINK_EM:AC005500.GENSCAN.20-2                          |   |     |

|    |        |          |  |   |
|----|--------|----------|--|---|
|    |        |          | CH22_FGENES.140_2  | 1.3   |
|    | 333619 | EOS33550 | CH22_880FG_219_3_LINK_EM:AC005500.GENSCAN.87-2                         |   |
|    |        |          | CH22_FGENES.219_3  | 1.3   |
| 5  | 335903 | EOS35834 | CH22_3280FG_635_11_LINK_EM:AC005500.GENSCAN.525-14                     |   |
|    |        |          | CH22_FGENES.635_11   | 1.3   |
|    | 326219 | EOS26150 | c17_hs gjl5867226[ref] gn 11 - 264008 264274 ex 3 5 CDSi 5.74 267 2847 |   |
|    |        |          | CH.17_hs gjl5867226  | 1.3   |
|    | 324456 | EOS24387 | AW500954   | EST cluster (not in UniGene)  |
|    | 316405 | EOS16336 | AA757900 Hs.202624   | ESTs  |
| 10 | 314361 | EOS14292 | AL038765 Hs.161304   | ESTs  |
|    | 329546 | EOS28477 | c_7_hs gjl5868487[ref] gn 1 - 17547 17722 ex 2 3 CDSi 9.96 176 3284    |   |
|    |        |          | CH.07_hs gjl5868487  | 1.3   |
|    | 335871 | EOS35802 | CH22_3246FG_629_19_LINK_EM:AC005500.GENSCAN.519-18                     |   |
|    |        |          | CH22_FGENES.629_19   | 1.3   |
| 15 | 303735 | EOS03666 | AA707750 Hs.202616   | ESTs; Weakly similar to cis-Golgi matrix protein GM130 [R.norvegicus]                             |
|    | 324048 | EOS23979 | AA378739   | EST cluster (not in UniGene)  |
|    | 326720 | EOS26651 | c20_hs gjl6552456[ref] gn 1 + 84525 84677 ex 5 7 CDSi 11.78 153 1031   |   |
|    |        |          | CH.20_hs gjl6552456  | 1.3   |
| 20 | 322309 | EOS22240 | AF086372   | EST cluster (not in UniGene)  |
|    | 322136 | EOS22067 | AF075083   | EST cluster (not in UniGene)  |
|    | 313460 | EOS13391 | AW028655 Hs.136033   | ESTs  |
|    | 306275 | EOS06206 | AA936312   | EST singleton (not in UniGene) with exon hit  |
|    | 321974 | EOS21905 | N76794   | EST cluster (not in UniGene)  |
| 25 | 327600 | EOS27531 | c_3_hs gjl6004462[ref] gn 1 - 2621 2862 ex 1 4 CDSi -4.01 242 1407     |   |
|    |        |          | CH.03_hs gjl6004462  | 1.3   |
|    | 329086 | EOS29017 | c_x_hs gjl5868604[ref] gn 1 - 35489 35588 ex 2 9 CDSi 2.55 100 719     |   |
|    |        |          | CH.X_hs gjl5868604   | 1.3   |
|    | 336919 | EOS36850 | CH22_4690FG_346_6_   | CH22_FGENES.346-6   |
| 30 | 302767 | EOS02698 | H94900 Hs.17882  | ESTs  |
|    | 334786 | EOS34717 | CH22_2098FG_432_11_LINK_EM:AC005500.GENSCAN.293-14                     |   |
|    |        |          | CH22_FGENES.432_11   | 1.3   |
|    | 302472 | EOS02403 | AA317451 Hs.241451   | SWI/SNF related; matrix associated; actin dependent regulator of chromatin; subfamily e; member 1 |
|    | 333033 | EOS32964 | CH22_259FG_68_8_LINK_EM:AC000097.GENSCAN.40-8                          |   |
| 35 |        |          | CH22_FGENES.68_8   | 1.3   |
|    | 330493 | EOS30424 | M27826 Hs.238380   | Human endogenous retroviral protease mRNA; complete cds   |
|    | 330506 | EOS30437 | M61906 Hs.6241   | phosphoinositide-3-kinase; regulatory subunit; polypeptide 1 (p85 alpha)                          |
|    | 313932 | EOS13863 | AI147601 Hs.154087   | ESTs  |
|    | 314394 | EOS14325 | AI380563 Hs.130816   | ESTs  |
|    | 323033 | EOS22964 | AI744284 Hs.221727   | ESTs  |
| 40 | 326431 | EOS26362 | c19_hs gjl5867371[ref] gn 1 + 15855 15971 ex 4 6 CDSi 7.79 117 1108    |   |
|    |        |          | CH.19_hs gjl5867371  | 1.3   |
|    | 335547 | EOS35478 | CH22_2902FG_576_8_LINK_EM:AC005500.GENSCAN.467-8                       |   |
|    |        |          | CH22_FGENES.576_8  | 1.3   |
| 45 | 300548 | EOS00479 | AI026836 Hs.114689   | ESTs  |
|    | 316504 | EOS16435 | AW135854 Hs.132458   | ESTs  |
|    | 335756 | EOS35687 | CH22_3123FG_604_5_LINK_EM:AC005500.GENSCAN.493-10                      |   |
|    |        |          | CH22_FGENES.604_5  | 1.3   |
|    | 301209 | EOS01140 | AI809912 Hs.159354   | ESTs  |
| 50 | 306610 | EOS06541 | AI000635   | EST singleton (not in UniGene) with exon hit  |
|    | 314439 | EOS14370 | AI539443 Hs.137447   | ESTs  |
|    | 315396 | EOS15327 | AW296107 Hs.152686   | ESTs  |
|    | 335914 | EOS35845 | CH22_3291FG_636_10_LINK_EM:AC005500.GENSCAN.526-10                     |   |
|    |        |          | CH22_FGENES.636_10   | 1.3   |
| 55 | 333734 | EOS33665 | CH22_1000FG_260_2_LINK_EM:AC005500.GENSCAN.119-7                       |   |
|    |        |          | CH22_FGENES.260_2  | 1.3   |
|    | 312370 | EOS12301 | AA744692 Hs.166539   | ESTs  |
|    | 304636 | EOS04567 | AA524031   | EST singleton (not in UniGene) with exon hit  |
|    | 323166 | EOS23097 | AA291001   | EST cluster (not in UniGene)  |
| 60 | 338702 | EOS38633 | CH22_7482FG_LINK_EM:AC005500.GENSCAN.480-1                             |   |
|    |        |          | CH22_EM:AC005500.GENSCAN.480-1   | 1.3   |
|    | 322331 | EOS22262 | AF086467   | EST cluster (not in UniGene)  |
|    | 318706 | EOS18637 | AI383593 Hs.159148   | ESTs  |
|    | 331186 | EOS31117 | T41159 Hs.8418   | ESTs  |
| 65 | 334764 | EOS34695 | CH22_2076FG_428_13_LINK_EM:AC005500.GENSCAN.289-13                     |   |
|    |        |          | CH22_FGENES.428_13   | 1.3   |
|    | 327565 | EOS27496 | c_3_hs gjl5867811[ref] gn 1 + 32516 32778 ex 2 3 CDSi 0.20 263 368     |   |
|    |        |          | CH.03_hs gjl5867811  | 1.3   |
|    | 335524 | EOS35455 | CH22_2879FG_572_4_LINK_EM:AC005500.GENSCAN.461-4                       |   |
|    |        |          | CH22_FGENES.572_4  | 1.3   |
| 70 | 308050 | EOS07981 | AI460004   | EST singleton (not in UniGene) with exon hit  |
|    | 334172 | EOS34103 | CH22_1452FG_349_5_LINK_EM:AC005500.GENSCAN.208-6                       |   |
|    |        |          | CH22_FGENES.349_5  | 1.3   |
|    | 315674 | EOS15605 | AA651923 Hs.191850   | ESTs  |
| 75 | 334876 | EOS34807 | CH22_2190FG_450_6_LINK_EM:AC005500.GENSCAN.339-6                       |   |
|    |        |          | CH22_FGENES.450_6  | 1.3   |
|    | 315606 | EOS15537 | AW298724 Hs.202639   | ESTs  |
|    | 338779 | EOS38710 | CH22_7610FG_LINK_EM:AC005500.GENSCAN.526-15                            |   |
|    |        |          | CH22_EM:AC005500.GENSCAN.526-15  | 1.3   |
| 80 | 333511 | EOS33442 | CH22_766FG_171_5_LINK_EM:AC005500.GENSCAN.51-5                         |   |
|    |        |          | CH22_FGENES.171_5  | 1.3   |
|    | 329254 | EOS29185 | c_x_hs gjl5868733[ref] gn 1 + 4133 4214 ex 1 2 CDSi -0.36 82 2833      |   |
|    |        |          | CH.X_hs gjl5868733   | 1.3   |
|    | 319510 | EOS19441 | W88633 Hs.254562   | ESTs  |
| 85 | 339418 | EOS39349 | CH22_8411FG_LINK_DJ579N16.GENSCAN.11-4                                 |   |
|    |        |          | CH22_DJ579N16.GENSCAN.11-4   | 1.3   |
|    | 321012 | EOS20943 | AA737314   | EST cluster (not in UniGene)  |

|    |        |          |  |  |
|----|--------|----------|--|--|
| 5  | 333217 | EOS33148 | CH22_454FG_104_9_LINK_EM:AC000097.GENSCAN.108-8                      | 1.3  |
|    |        |          | CH22_FGENES.104_9  |  |
|    | 338561 | EOS38492 | CH22_7294FG_LINK_EM:AC005500.GENSCAN.421-5                           | 1.3  |
|    |        |          | CH22_EM:AC005500.GENSCAN.421-5                                       |  |
|    | 335742 | EOS35673 | CH22_3105FG_601_13_LINK_EM:AC005500.GENSCAN.491-14                   | 1.3  |
| 10 |        |          | CH22_FGENES.601_13   |  |
|    | 334993 | EOS34924 | CH22_2314FG_469_14_LINK_EM:AC005500.GENSCAN.365-16                   | 1.3  |
|    |        |          | CH22_FGENES.469_14   |  |
|    | 323430 | EOS23361 | AW062479   | EST cluster (not in UniGene)   |
|    | 306069 | EOS06000 | AA906983   | EST singleton (not in UniGene) with exon hit                                     |
| 15 | 331681 | EOS31612 | W85712 Hs.119571   | collagen; type III; alpha 1 (Ehlers-Danlos syndrome type IV; autosomal dominant) |
|    | 337986 | EOS37917 | CH22_6441FG_LINK_EM:AC005500.GENSCAN.110-7                           | 1.3  |
|    |        |          | CH22_EM:AC005500.GENSCAN.110-7                                       |  |
|    | 313204 | EOS13135 | AI800518 Hs.118158   | ESTs   |
|    | 323189 | EOS23120 | AL121194 Hs.120589   | ESTs   |
| 20 | 318171 | EOS18102 | AA381202   | EST cluster (not in UniGene)   |
|    | 307156 | EOS07087 | AI186762   | EST singleton (not in UniGene) with exon hit                                     |
|    | 332713 | EOS32644 | AA349792 Hs.78489  | mutY (E. coli) homolog   |
|    | 312828 | EOS12759 | AI865455 Hs.211818   | ESTs; Moderately similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]  |
|    | 301127 | EOS01058 | AA758109 Hs.121072   | ESTs   |
| 25 | 311260 | EOS11191 | AI672509 Hs.196582   | ESTs   |
|    | 338364 | EOS38295 | CH22_7007FG_LINK_EM:AC005500.GENSCAN.323-7                           | 1.3  |
|    |        |          | CH22_EM:AC005500.GENSCAN.323-7                                       |  |
|    | 337904 | EOS37835 | CH22_6318FG_LINK_EM:AC005500.GENSCAN.56-17                           | 1.3  |
|    |        |          | CH22_EM:AC005500.GENSCAN.56-17                                       |  |
| 30 | 329347 | EOS29278 | c_x_hs gjl6456785[ref] gn 1 + 18433 18897 ex 4 4 CDSI 43.39 465 3718 | 1.3  |
|    |        |          | CH.X_hs gjl6456785   |  |
|    | 313329 | EOS13260 | AW293704 Hs.122658   | ESTs   |
|    | 314367 | EOS14298 | AA535749   | EST cluster (not in UniGene)   |
|    | 317098 | EOS17029 | AI123513 Hs.125456   | ESTs   |
| 35 | 306462 | EOS06393 | AA983397   | EST singleton (not in UniGene) with exon hit                                     |
|    | 301254 | EOS01185 | AI049624   | EST cluster (not in UniGene) with exon hit                                       |
|    | 335504 | EOS35435 | CH22_2856FG_571_15_LINK_EM:AC005500.GENSCAN.460-34                   | 1.3  |
|    |        |          | CH22_FGENES.571_15   |  |
|    | 334270 | EOS34201 | CH22_1559FG_368_2_LINK_EM:AC005500.GENSCAN.228-3                     | 1.3  |
| 40 |        |          | CH22_FGENES.368_2  |  |
|    | 334324 | EOS34255 | CH22_1616FG_375_1_LINK_EM:AC005500.GENSCAN.235-1                     | 1.3  |
|    |        |          | CH22_FGENES.375_1  |  |
|    | 304254 | EOS04185 | AA046273 Hs.111334   | ferritin; light polypeptide  |
|    | 305731 | EOS05662 | AA829363   | EST singleton (not in UniGene) with exon hit                                     |
| 45 | 323284 | EOS23215 | AA279381 Hs.190010   | ESTs   |
|    | 322007 | EOS21938 | AW410646 Hs.165739   | ESTs   |
|    | 334537 | EOS34468 | CH22_1839FG_403_2_LINK_EM:AC005500.GENSCAN.268-2                     | 1.3  |
|    |        |          | CH22_FGENES.403_2  |  |
|    | 302360 | EOS02291 | AJ010901 Hs.198267   | mucin 4; tracheobronchial  |
| 50 | 311641 | EOS11572 | AI948829 Hs.213786   | ESTs   |
|    | 324643 | EOS24574 | AI436356 Hs.130729   | ESTs   |
|    | 327554 | EOS27485 | c_3_hs gjl5867801[ref] gn 2 - 23092 23191 ex 2 6 CDSI 10.44 100 107  | 1.3  |
|    |        |          | CH.03_hs gjl5867801  |  |
|    | 312165 | EOS12096 | AW292139 Hs.115789   | ESTs   |
| 55 | 304679 | EOS04610 | AA548741   | EST singleton (not in UniGene) with exon hit                                     |
|    | 319564 | EOS19495 | AA026777 Hs.169732   | ESTs   |
|    | 310860 | EOS10791 | AW015920 Hs.161359   | ESTs   |
|    | 337161 | EOS37092 | CH22_5180FG_561_3_   | CH22_FGENES.561-3  |
|    | 311155 | EOS11086 | AI634410 Hs.197608   | EST  |
| 60 | 336846 | EOS36777 | CH22_4540FG_263_5_   | CH22_FGENES.263-5  |
|    | 310985 | EOS10916 | T51842   | EST cluster (not in UniGene)   |
|    | 329499 | EOS29430 | c10_p2 gjl3983518[gb]A gn 5 + 33463 33789 ex 1 1 CDSO 34.50 327 97   | 1.3  |
|    |        |          | CH.10_p2 gjl3983518  |  |
|    | 334924 | EOS34855 | CH22_2244FG_459_2_LINK_EM:AC005500.GENSCAN.351-2                     | 1.3  |
| 65 |        |          | CH22_FGENES.459_2  |  |
|    | 330861 | EOS30792 | AA084064 Hs.185747   | ESTs   |
|    | 324658 | EOS24589 | AI694767 Hs.129179   | ESTs   |
|    | 323362 | EOS23293 | AL135067 Hs.117182   | ESTs   |
|    | 330468 | EOS30399 | L10343 Hs.112341   | protease inhibitor 3; skin-derived (SKALP)                                       |
| 70 | 314198 | EOS14129 | AA897581 Hs.128773   | ESTs   |
|    | 339436 | EOS39367 | CH22_8431FG_LINK_DJ579N16.GENSCAN.19-1                               | 1.3  |
|    |        |          | CH22_DJ579N16.GENSCAN.19-1   |  |
|    | 312483 | EOS12414 | AI417526 Hs.184636   | ESTs   |
|    | 321505 | EOS21436 | H73183 Hs.129885   | ESTs   |
| 75 | 332254 | EOS32185 | N64702 Hs.194140   | ESTs   |
|    | 328253 | EOS28184 | c_6_hs gjl6381894[ref] gn 1 - 4411 4509 ex 1 5 CDSI 4.20 99 4561     | 1.3  |
|    |        |          | CH.06_hs gjl6381894  |  |
|    | 332357 | EOS32288 | W73417 Hs.103183   | EST  |
|    | 329017 | EOS28948 | c_x_hs gjl6682532[ref] gn 7 - 255591 255672 ex 3 3 CDSf 12.94 82 22  | 1.3  |
| 80 |        |          | CH.X_hs gjl6682532   |  |
|    | 337504 | EOS37435 | CH22_5739FG_803_2_   | CH22_FGENES.803-2  |
|    | 316625 | EOS16556 | AA780307 Hs.122156   | ESTs   |
|    | 335389 | EOS35320 | CH22_2739FG_545_1_LINK_EM:AC005500.GENSCAN.436-1                     | 1.3  |
|    |        |          | CH22_FGENES.545_1  |  |
| 85 | 310017 | EOS09948 | AI188739 Hs.148488   | ESTs   |
|    | 314354 | EOS14285 | AL037984 Hs.208982   | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]      |
|    | 324641 | EOS24572 | AI732515 Hs.189218   | ESTs   |
|    | 335207 | EOS35138 | CH22_2546FG_510_4_LINK_EM:AC005500.GENSCAN.402-3                     | 1.3  |
|    |        |          | CH22_FGENES.510_4  |  |
|    | 333673 | EOS33604 | CH22_934FG_246_5_LINK_EM:AC005500.GENSCAN.101-3                      | 1.3  |

|    |        |          |   |   |     |
|----|--------|----------|---|---|-----|
|    |        |          |   | CH22_FGENES.246_5   | 1.3 |
|    | 334370 | EOS34301 | CH22_1664FG_378_18_LINK_EM:AC005500.GENSCAN.240-1                     |   |     |
|    |        |          | CH22_FGENES.378_18  |   | 1.3 |
| 5  | 328690 | EOS28621 | c_7_hs_gi 5888001 ref  gn 7 - 571207 571274 ex 1 3 CDSi 3.34 68 4325  |   |     |
|    |        |          | CH.07_hs_gi 5888001   |   | 1.3 |
|    | 323208 | EOS23139 | AA203415 Hs.136200  | ESTs  | 1.3 |
|    | 307010 | EOS06941 | AI140014  | EST singleton (not in UniGene) with exon hit  | 1.3 |
|    | 316563 | EOS16494 | AI587083 Hs.200558  | ESTs; Weakly similar to !!!! ALU SUBFAMILY SP WARNING ENTRY !!!! [H.sapiens]                      | 1.3 |
| 10 | 312219 | EOS12150 | H73505 Hs.117874  | ESTs  | 1.3 |
|    | 319884 | EOS19815 | T73234  | EST cluster (not in UniGene)  | 1.3 |
|    | 334720 | EOS34651 | CH22_2030FG_421_31_LINK_EM:AC005500.GENSCAN.282-31                    |   |     |
|    |        |          | CH22_FGENES.421_31  |   | 1.3 |
|    | 335836 | EOS35767 | CH22_3210FG_621_3_LINK_EM:AC005500.GENSCAN.513-3                      |   |     |
| 15 |        |          | CH22_FGENES.621_3   |   | 1.3 |
|    | 305448 | EOS05379 | AA737894 Hs.29797   | ribosomal protein L10   | 1.3 |
|    | 314885 | EOS14816 | AI049878 Hs.133032  | ESTs  | 1.3 |
|    | 320130 | EOS20061 | AI820675 Hs.203804  | ESTs  | 1.3 |
|    | 310567 | EOS10498 | AI691065 Hs.155780  | ESTs  | 1.3 |
| 20 | 323898 | EOS23829 | AA347566  | EST cluster (not in UniGene)  | 1.3 |
|    | 336132 | EOS36063 | CH22_3522FG_703_2_LINK_DA59H18.GENSCAN.9-2                            |   |     |
|    |        |          | CH22_FGENES.703_2   |   | 1.3 |
|    | 337958 | EOS37889 | CH22_6403FG__LINK_EM:AC005500.GENSCAN.98-6                            |   |     |
| 25 |        |          | CH22_EM:AC005500.GENSCAN.98-6   |   | 1.3 |
|    | 305630 | EOS05561 | AA804508  | EST singleton (not in UniGene) with exon hit  | 1.3 |
|    | 334916 | EOS34847 | CH22_2235FG_457_7_LINK_EM:AC005500.GENSCAN.347-1                      |   |     |
|    |        |          | CH22_FGENES.457_7   |   | 1.3 |
|    | 333542 | EOS33473 | CH22_799FG_178_4_LINK_EM:AC005500.GENSCAN.59-4                        |   |     |
| 30 |        |          | CH22_FGENES.178_4   |   | 1.3 |
|    | 331151 | EOS1082  | R82331 Hs.164599  | ESTs  | 1.3 |
|    | 315095 | EOS15026 | AA831815 Hs.243788  | ESTs  | 1.3 |
|    | 331593 | EOS1524  | N72150 Hs.50193   | EST   | 1.3 |
|    | 323767 | EOS23698 | AI807408 Hs.166368  | ESTs  | 1.3 |
|    | 334561 | EOS34492 | CH22_1865FG_405_1_LINK_EM:AC005500.GENSCAN.270-5                      |   |     |
| 35 |        |          | CH22_FGENES.405_1   |   | 1.3 |
|    | 308191 | EOS08122 | AI538878  | EST singleton (not in UniGene) with exon hit  | 1.3 |
|    | 319571 | EOS19502 | N91399 Hs.220826  | ESTs  | 1.3 |
|    | 316200 | EOS16131 | AI914535 Hs.221377  | ESTs  | 1.3 |
|    | 305996 | EOS05927 | AA889338 Hs.163356  | EST   | 1.2 |
| 40 | 318055 | EOS17986 | AI249193 Hs.145945  | ESTs  | 1.2 |
|    | 315570 | EOS15501 | AI860360 Hs.160316  | ESTs  | 1.2 |
|    | 320792 | EOS20723 | AW236504 Hs.247020  | ESTs  | 1.2 |
|    | 331649 | EOS1580  | W20364 Hs.55412   | ESTs; Weakly similar to c29 [M.musculus]  | 1.2 |
|    | 303839 | EOS03770 | Z45939  | EST cluster (not in UniGene) with exon hit  | 1.2 |
| 45 | 324399 | EOS24330 | AA814768 Hs.21396   | ESTs  | 1.2 |
|    | 317172 | EOS17103 | AI741232 Hs.206744  | ESTs  | 1.2 |
|    | 312452 | EOS12383 | AI692643 Hs.172749  | ESTs  | 1.2 |
|    | 325482 | EOS25413 | c12_hs_gi 5866957 ref  gn 3 + 47957 48078 ex 5 7 CDSi 10.25 122 1896  |   |     |
|    |        |          | CH.12_hs_gi 5866957   |   | 1.2 |
| 50 | 311395 | EOS11326 | R23313  | EST cluster (not in UniGene)  | 1.2 |
|    | 336124 | EOS36055 | CH22_3513FG_701_9_LINK_DA59H18.GENSCAN.8-9                            |   |     |
|    |        |          | CH22_FGENES.701_9   |   | 1.2 |
|    | 320082 | EOS20013 | AA487678 Hs.189738  | ESTs  | 1.2 |
|    | 312168 | EOS12099 | T92251 Hs.198882  | ESTs  | 1.2 |
| 55 | 338000 | EOS37931 | CH22_6472FG__LINK_EM:AC005500.GENSCAN.119-5                           |   |     |
|    |        |          | CH22_EM:AC005500.GENSCAN.119-5  |   | 1.2 |
|    | 338852 | EOS38783 | CH22_7705FG__LINK_DJ246D7.GENSCAN.12-1                                |   |     |
|    |        |          | CH22_DJ246D7.GENSCAN.12-1   |   | 1.2 |
|    | 312090 | EOS12021 | N57692 Hs.118064  | ESTs  | 1.2 |
| 60 | 316480 | EOS16411 | AI749921 Hs.205377  | ESTs  | 1.2 |
|    | 333259 | EOS33190 | CH22_500FG_118_7_LINK_EM:AC005500.GENSCAN.2-7                         |   |     |
|    |        |          | CH22_FGENES.118_7   |   | 1.2 |
|    | 335211 | EOS35142 | CH22_2550FG_511_2_LINK_EM:AC005500.GENSCAN.403-2                      |   |     |
|    |        |          | CH22_FGENES.511_2   |   | 1.2 |
| 65 | 321950 | EOS21881 | AA594780 Hs.172318  | ESTs  | 1.2 |
|    | 337937 | EOS37868 | CH22_6370FG__LINK_EM:AC005500.GENSCAN.86-1                            |   |     |
|    |        |          | CH22_EM:AC005500.GENSCAN.86-1   |   | 1.2 |
|    | 316576 | EOS16507 | AI732114 Hs.193046  | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]                       | 1.2 |
|    | 322770 | EOS22701 | AA045796 Hs.159971  | SWI/SNF related; matrix associated; actin dependent regulator of chromatin; subfamily b; member 1 | 1.2 |
| 70 | 329369 | EOS29300 | c_x_hs_gi 5868842 ref  gn 1 - 121148 121516 ex 3 4 CDSi 8.50 369 3910 |   |     |
|    |        |          | CH.X_hs_gi 5868842  |   | 1.2 |
|    | 304183 | EOS04114 | H91161  | EST singleton (not in UniGene) with exon hit  | 1.2 |
|    | 339370 | EOS39301 | CH22_8343FG__LINK_BA232E17.GENSCAN.1-12                               |   |     |
|    |        |          | CH22_BA232E17.GENSCAN.1-12  |   | 1.2 |
| 75 | 303941 | EOS03872 | AW473878 Hs.156110  | Immunoglobulin kappa variable 1D-8  | 1.2 |
|    | 302245 | EOS02176 | H18835  | EST cluster (not in UniGene) with exon hit  | 1.2 |
|    | 335255 | EOS35186 | CH22_2597FG_517_2_LINK_EM:AC005500.GENSCAN.411-2                      |   |     |
|    |        |          | CH22_FGENES.517_2   |   | 1.2 |
|    | 316610 | EOS16541 | AW087973 Hs.126731  | ESTs  | 1.2 |
| 80 | 314915 | EOS14846 | AA573072 Hs.187748  | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]                       | 1.2 |
|    | 315426 | EOS15357 | AI391486 Hs.128171  | ESTs  | 1.2 |
|    | 334003 | EOS33934 | CH22_1281FG_310_28_LINK_EM:AC005500.GENSCAN.167-27                    |   |     |
|    |        |          | CH22_FGENES.310_28  |   | 1.2 |
|    | 304350 | EOS04281 | AA186871  | EST singleton (not in UniGene) with exon hit  | 1.2 |
| 85 | 325173 | EOS25104 | AI133215 Hs.144662  | ESTs; Moderately similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]                   | 1.2 |
|    | 312313 | EOS12244 | AW293341 Hs.122505  | ESTs  | 1.2 |
|    | 333366 | EOS33297 | CH22_612FG_142_3_LINK_EM:AC005500.GENSCAN.22-6                        |   |     |

|        |          |  |                   |     |
|--------|----------|--|-------------------|-----|
|        |          |  | CH22_FGENES.142_3 | 1.2 |
| 334970 | EOS34901 | CH22_2291FG_466_3_LINK_EM:AC005500.GENSCAN.361-2                                 |                   |     |
|        |          | CH22_FGENES.466_3  |                   | 1.2 |
| 5      | 338668   | EOS38599 CH22_7441FG_LINK_EM:AC005500.GENSCAN.465-1                              |                   |     |
|        |          | CH22_EM:AC005500.GENSCAN.465-1   |                   | 1.2 |
| 336502 | EOS36433 | CH22_3926FG_833_8_LINK_DJ579N16.GENSCAN.5-9                                      |                   |     |
|        |          | CH22_FGENES.833_8  |                   | 1.2 |
| 309438 | EOS09369 | AW102802 Hs.225787 ESTs; Moderately similar to hypothetical protein [H.sapiens]  |                   | 1.2 |
| 10     | 336194   | EOS36125 CH22_3591FG_717_20_LINK_DA59H18.GENSCAN.20-19                           |                   |     |
|        |          | CH22_FGENES.717_20   |                   | 1.2 |
| 336678 | EOS36609 | CH22_4156FG_43_6   |                   | 1.2 |
| 321401 | EOS21332 | W90406 Hs.35962 ESTs   |                   | 1.2 |
| 306026 | EOS05957 | AA902309 EST singleton (not in UniGene) with exon hit                            |                   | 1.2 |
| 15     | 336434   | EOS36365 CH22_3854FG_826_1_LINK_BA232E17.GENSCAN.8-1                             |                   |     |
|        |          | CH22_FGENES.826_1  |                   | 1.2 |
| 315257 | EOS15188 | AW157431 Hs.248941 ESTs  |                   | 1.2 |
| 328349 | EOS28280 | c.7_hs gjl5868383[ref] gn 7 - 260704 260804 ex 2 9 CDSi 4.37 101 621             |                   |     |
|        |          | CH.07_hs gjl5868383  |                   | 1.2 |
| 20     | 326112   | EOS26043 c.17_hs gjl5867192[ref] gn 1 + 2151 2725 ex 1 1 CDSi 54.87 575 1272     |                   |     |
|        |          | CH.17_hs gjl5867192  |                   | 1.2 |
| 333995 | EOS33926 | CH22_1272FG_310_19_LINK_EM:AC005500.GENSCAN.167-18                               |                   |     |
|        |          | CH22_FGENES.310_19   |                   | 1.2 |
| 323683 | EOS23614 | AI380045 Hs.225033 ESTs  |                   | 1.2 |
| 25     | 330143   | EOS30074 c.21_p2 gjl4210430[emb] gn 3 + 184737 184848 ex 4 4 CDSi 1.71 112 111   |                   |     |
|        |          | CH.21_p2 gjl4210430  |                   | 1.2 |
| 329789 | EOS29720 | c.14_p2 gjl6469354[emb] gn 2 - 118977 119036 ex 1 3 CDSi 1.19 60 1517            |                   |     |
|        |          | CH.14_p2 gjl6469354  |                   | 1.2 |
| 324397 | EOS24328 | AA307836 Hs.118758 ESTs; Weakly similar to RLF [H.sapiens]                       |                   | 1.2 |
| 30     | 308729   | EOS08660 AI799766 Hs.208627 EST  |                   | 1.2 |
| 323939 | EOS23870 | AW499632 Hs.115696 ESTs  |                   | 1.2 |
| 333444 | EOS33375 | CH22_694FG_153_1_LINK_EM:AC005500.GENSCAN.34-1                                   |                   |     |
|        |          | CH22_FGENES.153_1  |                   | 1.2 |
| 306302 | EOS06233 | AA937901 EST singleton (not in UniGene) with exon hit                            |                   | 1.2 |
| 35     | 313693   | EOS13624 AW469180 Hs.170651 ESTs   |                   | 1.2 |
| 316652 | EOS16583 | AA789249 EST cluster (not in UniGene)  |                   | 1.2 |
| 332325 | EOS32256 | T79428 Hs.191264 ESTs  |                   | 1.2 |
| 336235 | EOS36166 | CH22_3633FG_740_2_LINK_DA59H18.GENSCAN.44-2                                      |                   |     |
|        |          | CH22_FGENES.740_2  |                   | 1.2 |
| 319436 | EOS19367 | R02750 EST cluster (not in UniGene)  |                   | 1.2 |
| 40     | 312335   | EOS12266 AW043620 Hs.236993 ESTs   |                   | 1.2 |
| 322109 | EOS22040 | AI884327 Hs.244737 ESTs  |                   | 1.2 |
| 328466 | EOS28397 | c.7_hs gjl5868434[ref] gn 1 - 15643 15900 ex 1 2 CDSi 2.36 258 1608              |                   |     |
|        |          | CH.07_hs gjl5868434  |                   | 1.2 |
| 323244 | EOS23175 | T70731 EST cluster (not in UniGene)  |                   | 1.2 |
| 45     | 312510   | EOS12441 AA779907 Hs.117558 ESTs   |                   | 1.2 |
| 314853 | EOS14784 | AA729232 Hs.153279 ESTs  |                   | 1.2 |
| 336946 | EOS36877 | CH22_4731FG_355_2  |                   | 1.2 |
| 303874 | EOS03805 | AA258921 EST cluster (not in UniGene) with exon hit                              |                   | 1.2 |
| 50     | 312658   | EOS12589 AW30280 Hs.120936 ESTs  |                   | 1.2 |
| 308354 | EOS08285 | AI611044 EST singleton (not in UniGene) with exon hit                            |                   | 1.2 |
| 310073 | EOS10004 | AI335004 Hs.148558 ESTs  |                   | 1.2 |
| 324777 | EOS24708 | AA744046 Hs.133350 ESTs  |                   | 1.2 |
| 300897 | EOS00828 | AI890356 Hs.127804 ESTs  |                   | 1.2 |
| 55     | 308371   | EOS08302 AI620666 Hs.242510 EST  |                   | 1.2 |
| 306358 | EOS06289 | AA961821 EST singleton (not in UniGene) with exon hit                            |                   | 1.2 |
| 312295 | EOS12226 | AA578233 Hs.173863 ESTs  |                   | 1.2 |
| 319792 | EOS19723 | R20317 Hs.22968 ESTs   |                   | 1.2 |
| 338546 | EOS38477 | CH22_7267FG_LINK_EM:AC005500.GENSCAN.410-1                                       |                   |     |
|        |          | CH22_EM:AC005500.GENSCAN.410-1   |                   | 1.2 |
| 60     | 314546   | EOS14477 AW007211 Hs.186672 ESTs   |                   | 1.2 |
| 338494 | EOS38425 | CH22_7184FG_LINK_EM:AC005500.GENSCAN.385-5                                       |                   |     |
|        |          | CH22_EM:AC005500.GENSCAN.385-5   |                   | 1.2 |
| 331131 | EOS31062 | R54797 Hs.26238 EST; Weakly similar to reverse transcriptase homolog [H.sapiens] |                   | 1.2 |
| 65     | 309939   | EOS09870 AW419122 EST singleton (not in UniGene) with exon hit                   |                   | 1.2 |
| 332932 | EOS32863 | CH22_153FG_38_6_LINK_C20H12.GENSCAN.29-6   |                   |     |
|        |          | CH22_FGENES.38_6   |                   | 1.2 |
| 309653 | EOS09584 | AW196800 Hs.180842 ribosomal protein L13   |                   | 1.2 |
| 318647 | EOS18578 | AI526152 EST cluster (not in UniGene)  |                   | 1.2 |
| 70     | 304044   | EOS03975 T52479 Hs.252259 ribosomal protein S3                                   |                   | 1.2 |
| 330307 | EOS30238 | c.7_p2 gjl4877982[gb]A gn 2 + 107384 107559 ex 2 4 CDSi 9.96 176 4               |                   |     |
|        |          | CH.07_p2 gjl4877982  |                   | 1.2 |
| 314499 | EOS14430 | AL044570 Hs.147975 ESTs  |                   | 1.2 |
| 338053 | EOS37984 | CH22_6552FG_LINK_EM:AC005500.GENSCAN.158-1                                       |                   |     |
|        |          | CH22_EM:AC005500.GENSCAN.158-1   |                   | 1.2 |
| 75     | 332991   | EOS32922 CH22_215FG_56_4_LINK_EM:AC000097.GENSCAN.17-4                           |                   |     |
|        |          | CH22_FGENES.56_4   |                   | 1.2 |
| 306308 | EOS06239 | AA946870 EST singleton (not in UniGene) with exon hit                            |                   | 1.2 |
| 338120 | EOS38051 | CH22_6655FG_LINK_EM:AC005500.GENSCAN.195-1                                       |                   |     |
|        |          | CH22_EM:AC005500.GENSCAN.195-1   |                   | 1.2 |
| 80     | 313703   | EOS13634 AI161293 Hs.146862 ESTs; Weakly similar to KIAA0525 protein [H.sapiens] |                   | 1.2 |
| 330563 | EOS30494 | U50553 Hs.147916 DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 3                  |                   | 1.2 |
| 332886 | EOS32817 | CH22_106FG_33_7_LINK_C20H12.GENSCAN.22-9   |                   |     |
|        |          | CH22_FGENES.33_7   |                   | 1.2 |
| 303844 | EOS03775 | U94362 Hs.58589 glycogenin 2   |                   | 1.2 |
| 85     | 321755   | EOS21686 AI215881 Hs.144042 ESTs   |                   | 1.2 |
| 333532 | EOS33463 | CH22_789FG_175_19_LINK_EM:AC005500.GENSCAN.53-25                                 |                   |     |

|    |        |          |  |   |     |
|----|--------|----------|--|---|-----|
|    |        |          |  | CH22_FGENES.175_19  | 1.2 |
|    | 332863 | EOS32794 | CH22_81FG_28_3_LINK_C20H12.GENSCAN.18-3                                  |   |     |
|    |        |          |  | CH22_FGENES.28_3  | 1.2 |
| 5  | 333254 | EOS33185 | CH22_495FG_118_2_LINK_EM:AC005500.GENSCAN.2-2                            |   |     |
|    |        |          |  | CH22_FGENES.118_2   | 1.2 |
|    | 317459 | EOS17390 | AI367254 Hs.131248   | ESTs  | 1.2 |
|    | 315353 | EOS15284 | AW452608 Hs.129817   | ESTs  | 1.2 |
|    | 300732 | EOS00663 | AI369956 Hs.257891   | ESTs  | 1.2 |
|    | 303502 | EOS03433 | AA488528   | EST cluster (not in UniGene) with exon hit                                    | 1.2 |
| 10 | 333126 | EOS33057 | CH22_355FG_82_3_LINK_EM:AC000097.GENSCAN.66-10                           |   |     |
|    |        |          |  | CH22_FGENES.82_3  | 1.2 |
|    | 332929 | EOS32860 | CH22_150FG_38_3_LINK_C20H12.GENSCAN.29-3                                 |   |     |
|    |        |          |  | CH22_FGENES.38_3  | 1.2 |
| 15 | 329502 | EOS29433 | c10_p2 gij3983517 gbjU gn 1 + 75 338 ex 1 1 CDSi 46.82 264 100           |   |     |
|    |        |          |  | CH.10_p2 gij3983517   | 1.2 |
|    | 333408 | EOS33339 | CH22_657FG_145_6_LINK_EM:AC005500.GENSCAN.26-6                           |   |     |
|    |        |          |  | CH22_FGENES.145_6   | 1.2 |
|    | 315472 | EOS15403 | AA828850 Hs.165469   | ESTs  | 1.2 |
| 20 | 328290 | EOS28221 | c_7_hs gij5868363 refl gn 2 - 127366 127496 ex 1 5 CDSi 5.24 131 289     |   |     |
|    |        |          |  | CH.07_hs gij5868363   | 1.2 |
|    | 328662 | EOS28593 | c_7_hs gij6004473 refl gn 22 + 1184773 1184855 ex 7 8 CDSi 12.72 83 3916 |   |     |
|    |        |          |  | CH.07_hs gij6004473   | 1.2 |
|    | 319808 | EOS19739 | T58960   | EST cluster (not in UniGene)  | 1.2 |
|    | 303929 | EOS03860 | AW470753   | EST singleton (not in UniGene) with exon hit                                  | 1.2 |
| 25 | 315712 | EOS15643 | AI950133 Hs.120882   | ESTs; Moderately similar to !!! ALU SUBFAMILY J WARNING ENTRY !!! [H.sapiens] | 1.2 |
|    | 307391 | EOS07322 | AI225058   | EST singleton (not in UniGene) with exon hit                                  | 1.2 |
|    | 335499 | EOS35430 | CH22_2851FG_571_8_LINK_EM:AC005500.GENSCAN.460-28                        |   |     |
|    |        |          |  | CH22_FGENES.571_8   | 1.2 |
| 30 | 303792 | EOS03723 | C75094 Hs.199839   | ESTs; Highly similar to NG22 [H.sapiens]                                      | 1.2 |
|    | 327287 | EOS27218 | c_1_hs gij5867479 refl gn 1 - 62838 63024 ex 4 5 CDSi 11.66 187 1628     |   |     |
|    |        |          |  | CH.01_hs gij5867479   | 1.2 |
|    | 317713 | EOS17644 | AI733306 Hs.128071   | ESTs  | 1.2 |
|    | 330137 | EOS30068 | c21_p2 gij4210430 embj gn 1 - 21220 21377 ex 2 3 CDSi 1.89 158 104       |   |     |
| 35 |        |          |  | CH.21_p2 gij4210430   | 1.2 |
|    | 308157 | EOS08088 | AI510824 Hs.75968  | thymosin; beta 4; X chromosome  | 1.2 |
|    | 314452 | EOS14383 | AL042699 Hs.209222   | ESTs  | 1.2 |
|    | 308268 | EOS08199 | AI567509 Hs.172928   | collagen; type I; alpha 1   | 1.2 |
|    | 321467 | EOS21398 | X13075   | EST cluster (not in UniGene)  | 1.2 |
| 40 | 320993 | EOS20924 | AL050145 Hs.225986   | Homo sapiens mRNA; cDNA DKFZp586C2020 (from clone DKFZp586C2020)              | 1.2 |
|    | 336778 | EOS36709 | CH22_4367FG_159_4_   | CH22_FGENES.159-4   | 1.2 |
|    | 319827 | EOS19758 | T62778   | EST cluster (not in UniGene)  | 1.2 |
|    | 308249 | EOS08180 | AI560998   | EST singleton (not in UniGene) with exon hit                                  | 1.2 |
|    | 310094 | EOS10025 | AW450967 Hs.235240   | ESTs  | 1.2 |
| 45 | 336902 | EOS36833 | CH22_4655FG_331_2_   | CH22_FGENES.331-2   | 1.2 |
|    | 339044 | EOS38975 | CH22_7944FG_LINK_DA59H18.GENSCAN.27-5                                    |   |     |
|    |        |          |  | CH22_DA59H18.GENSCAN.27-5   | 1.2 |
|    | 336675 | EOS36606 | CH22_4153FG_43_3_  | CH22_FGENES.43-3  | 1.2 |
|    | 303563 | EOS03494 | AA367699 Hs.118787   | transforming growth factor; beta-induced; 68kD                                | 1.2 |
| 50 | 330673 | EOS30604 | D57823 Hs.92962  | Sec23 (S. cerevisiae) homolog A   | 1.2 |
|    | 311814 | EOS11745 | AW377113 Hs.119640   | ESTs; Moderately similar to zinc finger protein [H.sapiens]                   | 1.2 |
|    | 335481 | EOS35412 | CH22_2833FG_570_10_LINK_EM:AC005500.GENSCAN.460-4                        |   |     |
|    |        |          |  | CH22_FGENES.570_10  | 1.2 |
|    | 314775 | EOS14706 | AI149880 Hs.188809   | ESTs  | 1.2 |
| 55 | 324961 | EOS24892 | AA613792   | EST cluster (not in UniGene)  | 1.2 |
|    | 313458 | EOS13389 | AA007259 Hs.255853   | ESTs  | 1.2 |
|    | 307074 | EOS07005 | AI150989   | EST singleton (not in UniGene) with exon hit                                  | 1.2 |
|    | 337964 | EOS37895 | CH22_6410FG_LINK_EM:AC005500.GENSCAN.100-9                               |   |     |
|    |        |          |  | CH22_EM:AC005500.GENSCAN.100-9  | 1.2 |
| 60 | 326519 | EOS26450 | c19_hs gij5867439 refl gn 4 + 166004 166243 ex 4 5 CDSi 4.49 240 2534    |   |     |
|    |        |          |  | CH.19_hs gij5867439   | 1.2 |
|    | 337366 | EOS37297 | CH22_5551FG_736_1_   | CH22_FGENES.736-1   | 1.2 |
|    | 322340 | EOS22271 | AF088076   | EST cluster (not in UniGene)  | 1.2 |
|    | 307954 | EOS07885 | AI419692   | EST singleton (not in UniGene) with exon hit                                  | 1.2 |
| 65 | 328615 | EOS28546 | c_7_hs gij5868239 refl gn 2 + 35214 35347 ex 3 4 CDSi 11.49 134 3651     |   |     |
|    |        |          |  | CH.07_hs gij5868239   | 1.2 |
|    | 317787 | EOS17718 | AW339612 Hs.249364   | ESTs  | 1.2 |
|    | 335288 | EOS35219 | CH22_2630FG_527_1_LINK_EM:AC005500.GENSCAN.421-1                         |   |     |
|    |        |          |  | CH22_FGENES.527_1   | 1.2 |
| 70 | 323175 | EOS23106 | AI827137 Hs.184023   | ESTs  | 1.2 |
|    | 330893 | EOS30824 | AA149620 Hs.71999  | ESTs  | 1.2 |
|    | 306810 | EOS06741 | AI057294   | EST singleton (not in UniGene) with exon hit                                  | 1.2 |
|    | 338239 | EOS38170 | CH22_6833FG_LINK_EM:AC005500.GENSCAN.264-5                               |   |     |
|    |        |          |  | CH22_EM:AC005500.GENSCAN.264-5  | 1.2 |
| 75 | 332347 | EOS32278 | W60326 Hs.221716   | ESTs  | 1.2 |
|    | 309782 | EOS09713 | AW275156 Hs.156110   | Immunoglobulin kappa variable 1D-8  | 1.2 |
|    | 322518 | EOS22449 | AI133446   | EST cluster (not in UniGene)  | 1.2 |
|    | 301187 | EOS01118 | AA806542   | EST cluster (not in UniGene) with exon hit                                    | 1.2 |
|    | 312129 | EOS12060 | AW300867   | EST cluster (not in UniGene)  | 1.2 |
| 80 | 334714 | EOS34645 | CH22_2024FG_421_25_LINK_EM:AC005500.GENSCAN.282-25                       |   |     |
|    |        |          |  | CH22_FGENES.421_25  | 1.2 |
|    | 316586 | EOS16517 | AI205077 Hs.144689   | ESTs  | 1.2 |
|    | 320488 | EOS20419 | R31386   | EST cluster (not in UniGene)  | 1.2 |
|    | 327458 | EOS27389 | c_2_hs gij6004455 refl gn 3 + 173257 173378 ex 5 7 CDSi 4.03 122 1184    |   |     |
| 85 |        |          |  | CH.02_hs gij6004455   | 1.2 |
|    | 336707 | EOS36638 | CH22_4212FG_64_3_  | CH22_FGENES.64-3  | 1.2 |
|    | 313561 | EOS13492 | AA040155   | EST cluster (not in UniGene)  | 1.2 |

|    |        |          |                                    |   |  |     |
|----|--------|----------|------------------------------------|---|--|-----|
|    | 330906 | EOS30837 | AA169498                           | Hs.72804                                    | ESTs   | 1.2 |
|    | 330987 | EOS30918 | H40988                             | Hs.131965                                   | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]      | 1.2 |
|    | 325041 | EOS24972 | AI809182                           | Hs.130907                                   | ESTs   | 1.2 |
| 5  | 313225 | EOS13156 | AA502384                           | Hs.151529                                   | ESTs   | 1.2 |
|    | 305295 | EOS05226 | AA687131                           |   | EST singleton (not in UniGene) with exon hit                                     | 1.2 |
|    | 306896 | EOS06827 | AI093383                           |   | EST singleton (not in UniGene) with exon hit                                     | 1.2 |
|    | 326981 | EOS26912 | c21_hs gjl[6588016]refl gn 3       | + 105091 106038 ex 1 1 CDSi 122.69 948 567  | CH.21_hs gjl[6588016]  | 1.2 |
|    | 332225 | EOS32156 | N33213                             | Hs.100425                                   | ESTs   | 1.2 |
| 10 | 318802 | EOS18733 | R19443                             | Hs.92414                                    | ESTs   | 1.2 |
|    | 318413 | EOS18344 | AI138592                           | Hs.144936                                   | ESTs   | 1.2 |
|    | 312292 | EOS12223 | AW451893                           | Hs.151124                                   | ESTs   | 1.2 |
|    | 323763 | EOS23684 | AA327102                           |   | EST cluster (not in UniGene)   | 1.2 |
| 15 | 313582 | EOS13513 | AW207684                           | Hs.13583                                    | ESTs   | 1.2 |
|    | 317836 | EOS17767 | AA983913                           | Hs.128929                                   | ESTs   | 1.2 |
|    | 332868 | EOS32799 | CH22_86FG_28_8_LINK_C20H12         | GENSCAN.18-8                                | CH22_FGENES.28_8   | 1.2 |
|    | 336924 | EOS36855 | CH22_4699FG_347_9_                 |   | CH22_FGENES.347-9  | 1.2 |
| 20 | 327791 | EOS27722 | c_5_hs gjl[5867977]refl gn 1       | + 22491 22610 ex 6 7 CDSi 11.29 120 658     | CH.05_hs gjl[5867977]  | 1.2 |
|    | 330717 | EOS30648 | AA233926                           | Hs.23635                                    | ESTs   | 1.2 |
|    | 322944 | EOS22875 | AA112573                           |   | EST cluster (not in UniGene)   | 1.2 |
|    | 312108 | EOS12039 | T82331                             | Hs.127453                                   | ESTs   | 1.2 |
| 25 | 332570 | EOS32501 | AA401376                           | Hs.26176                                    | ESTs   | 1.2 |
|    | 330880 | EOS30811 | AA132420                           | Hs.53542                                    | KIAA0986 protein   | 1.2 |
|    | 310341 | EOS10272 | AW302773                           |   | EST cluster (not in UniGene)   | 1.2 |
|    | 334012 | EOS33943 | CH22_1290FG_313_3_LINK_EM:AC005500 | GENSCAN.169-3                               | CH22_FGENES.313_3  | 1.2 |
| 30 | 318230 | EOS18161 | AA558125                           |   | EST cluster (not in UniGene)   | 1.2 |
|    | 336071 | EOS36002 | CH22_3457FG_685_3_LINK_DJ32110     | GENSCAN.21-6                                | CH22_FGENES.685_3  | 1.2 |
|    | 338510 | EOS38441 | CH22_7208FG__LINK_EM:AC005500      | GENSCAN.391-22                              | CH22_EM:AC005500.GENSCAN.391-22  | 1.2 |
| 35 | 334487 | EOS34418 | CH22_1786FG_395_9_LINK_EM:AC005500 | GENSCAN.258-10                              | CH22_FGENES.395_9  | 1.2 |
|    | 320661 | EOS20592 | AA864846                           |   | EST cluster (not in UniGene)   | 1.2 |
|    | 335200 | EOS35131 | CH22_2538FG_508_9_LINK_EM:AC005500 | GENSCAN.401-9                               | CH22_FGENES.508_9  | 1.2 |
| 40 | 333582 | EOS33513 | CH22_842FG_201_2_LINK_EM:AC005500  | GENSCAN.72-3                                | CH22_FGENES.201_2  | 1.2 |
|    | 320789 | EOS20720 | R78712                             |   | EST cluster (not in UniGene)   | 1.2 |
|    | 321185 | EOS21116 | H51659                             | Hs.189854                                   | ESTs   | 1.2 |
|    | 337740 | EOS37671 | CH22_6085FG__LINK_EM:AC000097      | GENSCAN.100-6                               | CH22_EM:AC000097.GENSCAN.100-6   | 1.2 |
| 45 | 315064 | EOS14995 | AA775208                           | Hs.136423                                   | ESTs   | 1.2 |
|    | 334883 | EOS34814 | CH22_2197FG_451_6_LINK_EM:AC005500 | GENSCAN.340-6                               | CH22_FGENES.451_6  | 1.2 |
| 50 | 331825 | EOS31756 | AA411144                           | Hs.104768                                   | ESTs   | 1.2 |
|    | 319141 | EOS19072 | F12377                             |   | EST cluster (not in UniGene)   | 1.1 |
|    | 333682 | EOS33613 | CH22_944FG_247_10_LINK_EM:AC005500 | GENSCAN.102-10                              | CH22_FGENES.247_10   | 1.1 |
|    | 336140 | EOS36071 | CH22_3530FG_705_2_LINK_DA59H18     | GENSCAN.10-2                                | CH22_FGENES.705_2  | 1.1 |
| 55 | 320727 | EOS20658 | U96044                             |   | EST cluster (not in UniGene)   | 1.1 |
|    | 323947 | EOS23878 | AA649842                           | Hs.186667                                   | ESTs   | 1.1 |
|    | 324746 | EOS24677 | AA603367                           | Hs.222294                                   | ESTs   | 1.1 |
|    | 306744 | EOS06675 | AI031882                           |   | EST singleton (not in UniGene) with exon hit                                     | 1.1 |
|    | 326517 | EOS26448 | c19_hs gjl[5867439]refl gn 1       | + 44732 46356 ex 6 6 CDSi 148.22 1625 2512  | CH.19_hs gjl[5867439]  | 1.1 |
| 60 | 333597 | EOS33528 | CH22_858FG_211_5_LINK_EM:AC005500  | GENSCAN.79-5                                | CH22_FGENES.211_5  | 1.1 |
|    | 330135 | EOS30066 | c21_p2 gjl[4456470]embj gn 2       | - 121583 121885 ex 2 2 CDSf 18.67 303 102   | CH.21_p2 gjl[4456470]  | 1.1 |
| 65 | 315118 | EOS15049 | AA564921                           | Hs.143899                                   | ESTs   | 1.1 |
|    | 302893 | EOS02824 | AL117539                           | Hs.173515                                   | Homo sapiens mRNA; cDNA DKFZp586H021 (from clone DKFZp586H021)                   | 1.1 |
|    | 337169 | EOS37100 | CH22_5189FG_563_1_                 |   | CH22_FGENES.563-1  | 1.1 |
|    | 336121 | EOS36052 | CH22_3510FG_701_6_LINK_DA59H18     | GENSCAN.8-6                                 | CH22_FGENES.701_6  | 1.1 |
| 70 | 323332 | EOS23263 | AI829520                           | Hs.227513                                   | ESTs   | 1.1 |
|    | 320911 | EOS20842 | AI056872                           | Hs.133386                                   | ESTs   | 1.1 |
|    | 327990 | EOS27921 | c_6_hs gjl[5868218]refl gn 2       | - 36225 36503 ex 1 2 CDSi 16.35 279 1419    | CH.06_hs gjl[5868218]  | 1.1 |
| 75 | 320425 | EOS20356 | C14069                             | Hs.201627                                   | ESTs; Moderately similar to !!!! ALU SUBFAMILY SQ WARNING ENTRY !!!! [H.sapiens] | 1.1 |
|    | 327075 | EOS27006 | c21_hs gjl[6531965]refl gn 58      | + 4041318 4041431 ex 4 4 CDSi 1.79 114 1285 | CH.21_hs gjl[6531965]  | 1.1 |
|    | 314384 | EOS14315 | AA535840                           | Hs.162203                                   | ESTs; Weakly similar to alternatively spliced product using exon 13A [H.sapiens] | 1.1 |
|    | 338716 | EOS38647 | CH22_7502FG__LINK_EM:AC005500      | GENSCAN.488-9                               | CH22_EM:AC005500.GENSCAN.488-9   | 1.1 |
| 80 | 330886 | EOS30817 | AA135606                           | Hs.189384                                   | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]      | 1.1 |
|    | 327331 | EOS27262 | c_1_hs gjl[5867516]refl gn 4       | - 55606 55737 ex 2 6 CDSi 7.01 132 2349     | CH.01_hs gjl[5867516]  | 1.1 |
|    | 326714 | EOS26645 | c20_hs gjl[5867595]refl gn 2       | + 124490 124568 ex 5 6 CDSi 0.11 79 1020    | CH.20_hs gjl[5867595]  | 1.1 |
| 85 | 316734 | EOS16665 | AW080237                           | Hs.252884                                   | ESTs   | 1.1 |
|    | 311660 | EOS11591 | AI978583                           | Hs.232161                                   | ESTs   | 1.1 |
|    | 312757 | EOS12688 | AI285970                           | Hs.183817                                   | ESTs   | 1.1 |

|    |        |          |  |                   |  |     |
|----|--------|----------|--|-------------------|--|-----|
|    | 331686 | EOS31617 | W88502   | Hs.182258         | ESTs   | 1.1 |
|    | 337840 | EOS37771 | CH22_6223FG_LINK_EM:AC005500.GENSCAN.26-9                            |                   |  | 1.1 |
|    |        |          | CH22_EM:AC005500.GENSCAN.26-9  |                   |  | 1.1 |
| 5  | 332093 | EOS32024 | AA608794   | Hs.112592         | ESTs   | 1.1 |
|    | 319595 | EOS19526 | H81361   | Hs.194485         | ESTs   | 1.1 |
|    | 315990 | EOS15921 | AI800041   | Hs.190555         | ESTs   | 1.1 |
|    | 322438 | EOS22369 | W44531   | Hs.167851         | ESTs   | 1.1 |
|    | 332965 | EOS32896 | CH22_189FG_50_3_LINK_EM:AC000097.GENSCAN.3-5                         |                   |  | 1.1 |
|    |        |          | CH22_FGENES.50_3   |                   |  | 1.1 |
| 10 | 337182 | EOS37113 | CH22_5204FG_570_2_   | CH22_FGENES.570-2 |  | 1.1 |
|    | 334948 | EOS34879 | CH22_2269FG_465_15_LINK_EM:AC005500.GENSCAN.359-13                   |                   |  | 1.1 |
|    |        |          | CH22_FGENES.465_15   |                   |  | 1.1 |
|    | 325864 | EOS25795 | c16_hs gjl5867069 refl gn 2 - 110834 110904 ex 3 3 CDSf 9.76 71 457  |                   |  | 1.1 |
|    |        |          | CH.16_hs gjl5867069  |                   |  | 1.1 |
| 15 | 337760 | EOS37691 | CH22_6110FG_LINK_EM:AC000097.GENSCAN.116-8                           |                   |  | 1.1 |
|    |        |          | CH22_EM:AC000097.GENSCAN.116-8                                       |                   |  | 1.1 |
|    | 315422 | EOS15353 | AW135357   | Hs.192374         | ESTs   | 1.1 |
|    | 338889 | EOS38820 | CH22_7746FG_LINK_DJ32110.GENSCAN.7-1                                 |                   |  | 1.1 |
|    |        |          | CH22_DJ32110.GENSCAN.7-1   |                   |  | 1.1 |
| 20 | 332961 | EOS32892 | CH22_185FG_48_18_LINK_EM:AC000097.GENSCAN.2-14                       |                   |  | 1.1 |
|    |        |          | CH22_FGENES.48_18  |                   |  | 1.1 |
|    | 314703 | EOS14634 | AI791249   |                   | EST cluster (not in UniGene)   | 1.1 |
|    | 317791 | EOS17722 | AI801500   | Hs.128457         | ESTs   | 1.1 |
|    | 333680 | EOS33611 | CH22_942FG_247_7_LINK_EM:AC005500.GENSCAN.102-7                      |                   |  | 1.1 |
| 25 |        |          | CH22_FGENES.247_7  |                   |  | 1.1 |
|    | 322419 | EOS22350 | AA248987   | Hs.14084          | ESTs; Highly similar to zinc RING finger protein SAG [M.musculus]            | 1.1 |
|    | 338124 | EOS38055 | CH22_6661FG_LINK_EM:AC005500.GENSCAN.196-2                           |                   |  | 1.1 |
|    |        |          | CH22_EM:AC005500.GENSCAN.196-2                                       |                   |  | 1.1 |
| 30 | 308884 | EOS08815 | AI833131   | Hs.179100         | ESTs   | 1.1 |
|    | 333349 | EOS33280 | CH22_595FG_140_3_LINK_EM:AC005500.GENSCAN.20-3                       |                   |  | 1.1 |
|    |        |          | CH22_FGENES.140_3  |                   |  | 1.1 |
|    | 313150 | EOS13081 | AA824410   | Hs.165003         | ESTs   | 1.1 |
|    | 339208 | EOS39139 | CH22_8146FG_LINK_FF113D11.GENSCAN.6-3                                |                   |  | 1.1 |
|    |        |          | CH22_FF113D11.GENSCAN.6-3  |                   |  | 1.1 |
| 35 | 335653 | EOS35584 | CH22_3013FG_590_4_LINK_EM:AC005500.GENSCAN.484-4                     |                   |  | 1.1 |
|    |        |          | CH22_FGENES.590_4  |                   |  | 1.1 |
|    | 319524 | EOS19455 | AA682865   | Hs.194441         | ESTs   | 1.1 |
|    | 301576 | EOS01507 | AI682905   | Hs.146875         | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]  | 1.1 |
|    | 317598 | EOS17529 | AW206035   | Hs.192123         | ESTs   | 1.1 |
| 40 | 333473 | EOS33404 | CH22_724FG_162_3_LINK_EM:AC005500.GENSCAN.42-10                      |                   |  | 1.1 |
|    |        |          | CH22_FGENES.162_3  |                   |  | 1.1 |
|    | 333949 | EOS33880 | CH22_1225FG_303_5_LINK_EM:AC005500.GENSCAN.162-9                     |                   |  | 1.1 |
|    |        |          | CH22_FGENES.303_5  |                   |  | 1.1 |
| 45 | 339256 | EOS39187 | CH22_8207FG_LINK_BA354112.GENSCAN.7-11                               |                   |  | 1.1 |
|    |        |          | CH22_BA354112.GENSCAN.7-11   |                   |  | 1.1 |
|    | 332884 | EOS32815 | CH22_104FG_33_5_LINK_C20H12.GENSCAN.22-7                             |                   |  | 1.1 |
|    |        |          | CH22_FGENES.33_5   |                   |  | 1.1 |
|    | 314660 | EOS14591 | AA436007   | Hs.188780         | ESTs   | 1.1 |
| 50 | 333220 | EOS33151 | CH22_457FG_104_12_LINK_EM:AC000097.GENSCAN.108-11                    |                   |  | 1.1 |
|    |        |          | CH22_FGENES.104_12   |                   |  | 1.1 |
|    | 308106 | EOS08037 | AI476803   |                   | EST singleton (not in UniGene) with exon hit                                 | 1.1 |
|    | 320709 | EOS20640 | AA456660   | Hs.154165         | ESTs   | 1.1 |
|    | 307612 | EOS07543 | AI290787   |                   | EST singleton (not in UniGene) with exon hit                                 | 1.1 |
| 55 | 330286 | EOS30217 | c_5_p2 gjl6671913 gb A gn 2 - 31050 31171 ex 2 7 CDSi 8.84 122 791   |                   |  | 1.1 |
|    |        |          | CH.05_p2 gjl6671913  |                   |  | 1.1 |
|    | 304495 | EOS04426 | AA446448   |                   | EST singleton (not in UniGene) with exon hit                                 | 1.1 |
|    | 310583 | EOS10514 | AW205632   | Hs.211198         | ESTs   | 1.1 |
|    | 332896 | EOS32827 | CH22_117FG_35_10_LINK_C20H12.GENSCAN.24-9                            |                   |  | 1.1 |
|    |        |          | CH22_FGENES.35_10  |                   |  | 1.1 |
| 60 | 337602 | EOS37533 | CH22_5895FG_LINK_C20H12.GENSCAN.15-1                                 |                   |  | 1.1 |
|    |        |          | CH22_C20H12.GENSCAN.15-1   |                   |  | 1.1 |
|    | 307626 | EOS07557 | AI300035   |                   | EST singleton (not in UniGene) with exon hit                                 | 1.1 |
|    | 334696 | EOS34627 | CH22_2006FG_421_5_LINK_EM:AC005500.GENSCAN.282-5                     |                   |  | 1.1 |
|    |        |          | CH22_FGENES.421_5  |                   |  | 1.1 |
| 65 | 318652 | EOS18583 | T53259   |                   | EST cluster (not in UniGene)   | 1.1 |
|    | 337844 | EOS37775 | CH22_6229FG_LINK_EM:AC005500.GENSCAN.30-9                            |                   |  | 1.1 |
|    |        |          | CH22_EM:AC005500.GENSCAN.30-9  |                   |  | 1.1 |
|    | 334823 | EOS34754 | CH22_2137FG_437_5_LINK_EM:AC005500.GENSCAN.301-7                     |                   |  | 1.1 |
|    |        |          | CH22_FGENES.437_5  |                   |  | 1.1 |
| 70 | 333928 | EOS33859 | CH22_1201FG_299_2_LINK_EM:AC005500.GENSCAN.158-5                     |                   |  | 1.1 |
|    |        |          | CH22_FGENES.299_2  |                   |  | 1.1 |
|    | 337503 | EOS37434 | CH22_5738FG_803_1_   | CH22_FGENES.803-1 |  | 1.1 |
|    | 323044 | EOS22975 | AA148725   | Hs.154190         | ESTs   | 1.1 |
| 75 | 329164 | EOS29095 | c_x_hs gjl5868691 refl gn 1 + 62305 62517 ex 2 2 CDSi 17.51 213 1868 |                   |  | 1.1 |
|    |        |          | CH.X_hs gjl5868691   |                   |  | 1.1 |
|    | 335468 | EOS35399 | CH22_2819FG_567_4_LINK_EM:AC005500.GENSCAN.454-12                    |                   |  | 1.1 |
|    |        |          | CH22_FGENES.567_4  |                   |  | 1.1 |
|    | 338962 | EOS38893 | CH22_7838FG_LINK_DJ32110.GENSCAN.23-39                               |                   |  | 1.1 |
|    |        |          | CH22_DJ32110.GENSCAN.23-39   |                   |  | 1.1 |
| 80 | 323570 | EOS23501 | AL038623   | Hs.208752         | ESTs; Weakly similar to !!!! ALU SUBFAMILY SX WARNING ENTRY !!!! [H.sapiens] | 1.1 |
|    | 333568 | EOS33499 | CH22_826FG_185_1_LINK_EM:AC005500.GENSCAN.64-1                       |                   |  | 1.1 |
|    |        |          | CH22_FGENES.185_1  |                   |  | 1.1 |
|    | 331865 | EOS31796 | AA425756   | Hs.98445          | ESTs   | 1.1 |
| 85 | 336246 | EOS36177 | CH22_3644FG_746_5_LINK_DA59H18.GENSCAN.48-4                          |                   |  | 1.1 |
|    |        |          | CH22_FGENES.746_5  |                   |  | 1.1 |
|    | 337238 | EOS37169 | CH22_5343FG_641_3_   | CH22_FGENES.641-3 |  | 1.1 |

|    |        |          |   |                 |   |     |
|----|--------|----------|---|-----------------|---|-----|
| 5  | 305089 | EOS05020 | AA642622  |                 | EST singleton (not in UniGene) with exon hit  | 1.1 |
|    | 300097 | EOS00028 | AI916973  | Hs.213603       | ESTs  | 1.1 |
|    | 313134 | EOS13065 | N63406  | Hs.258697       | ESTs  | 1.1 |
|    | 337452 | EOS37383 | CH22_5665FG_775_1_                                      |                 | CH22_FGENES.775-1   | 1.1 |
|    | 325433 | EOS25364 | c12_hs_gij5866936[ref] gn 4 - 480706 480826 ex 3 4 CDSi | 1.99 121 818    |   | 1.1 |
| 10 | 335999 | EOS35930 | CH22_3380FG_657_1_LINK_DJ246D7                          |                 | GENSCAN.11-1  | 1.1 |
|    | 333580 | EOS33511 | CH22_840FG_199_2_LINK_EM:AC005500                       |                 | GENSCAN.71-2  | 1.1 |
|    | 336836 | EOS36767 | CH22_4512FG_247_11_                                     |                 | CH22_FGENES.199_2   | 1.1 |
|    | 334677 | EOS34608 | CH22_1986FG_418_30_LINK_EM:AC005500                     |                 | GENSCAN.279-31  | 1.1 |
|    | 329062 | EOS28993 | c_x_hs_gij5868590[ref] gn 3 - 58977 59094 ex 4 11 CDSi  | -6.19 118 627   |   | 1.1 |
| 15 | 333671 | EOS33602 | CH22_932FG_245_5_LINK_EM:AC005500                       |                 | GENSCAN.100-12  | 1.1 |
|    | 304941 | EOS04872 | AA612612  |                 | EST singleton (not in UniGene) with exon hit  | 1.1 |
|    | 315772 | EOS15703 | AW515373  | Hs.158893       | ESTs  | 1.1 |
|    | 301281 | EOS01212 | AA843986  | Hs.190586       | ESTs  | 1.1 |
|    | 333520 | EOS33451 | CH22_777FG_174_3_LINK_EM:AC005500                       |                 | GENSCAN.53-6  | 1.1 |
| 20 | 315203 | EOS15134 | AI559820  | Hs.199438       | ESTs  | 1.1 |
|    | 315927 | EOS15858 | AW025517  | Hs.133250       | ESTs  | 1.1 |
|    | 317161 | EOS17092 | AA972165  | Hs.150308       | ESTs  | 1.1 |
|    | 337692 | EOS37623 | CH22_6028FG_LINK_EM:AC000097                            |                 | GENSCAN.78-12   | 1.1 |
|    | 331472 | EOS31403 | N24830  |                 | CH22_EM:AC000097.GENSCAN.78-12<br>yx70a02.s1 Soares melanocyte 2NbHM Homo sapiens cDNA clone IMAGE:267050 3' similar to gb U87912 HUMALNE562 Human carcinoma cell-derived Alu RNA transcript, (rRNA); contains Alu repetitive element; mRNA sequence. | 1.1 |
| 30 | 336439 | EOS36370 | CH22_3859FG_827_4_LINK_DJ579N16                         |                 | GENSCAN.1-3   | 1.1 |
|    | 326882 | EOS26813 | c20_hs_gij6682509[ref] gn 2 - 167988 168179 ex 4 4 CDSf | 18.69 192 2238  |   | 1.1 |
|    | 336977 | EOS36908 | CH22_4793FG_380_9_                                      |                 | CH22_FGENES.827_4   | 1.1 |
|    | 333983 | EOS33914 | CH22_1260FG_310_7_LINK_EM:AC005500                      |                 | GENSCAN.167-5   | 1.1 |
|    | 328878 | EOS28809 | c_7_hs_gij6552423[ref] gn 1 + 105580 105774 ex 6 7 CDSi | 2.91 195 6195   |   | 1.1 |
| 40 | 330415 | EOS30346 | D83777  | Hs.75137        | KIAA0193 gene product   | 1.1 |
|    | 324824 | EOS24755 | AI826999  | Hs.224624       | ESTs  | 1.1 |
|    | 325815 | EOS25746 | c14_hs_gij6682483[ref] gn 1 - 129273 130754 ex 1 1 CDSs | 11.82 1482 2225 |   | 1.1 |
|    | 300463 | EOS00394 | N52510  | Hs.186470       | ESTs  | 1.1 |
|    | 335708 | EOS35639 | CH22_3069FG_599_8_LINK_EM:AC005500                      |                 | GENSCAN.490-11  | 1.1 |
| 45 | 324575 | EOS24506 | AW502257  |                 | EST cluster (not in UniGene)  | 1.1 |
|    | 337951 | EOS37882 | CH22_6391FG_LINK_EM:AC005500                            |                 | GENSCAN.94-1  | 1.1 |
|    | 335935 | EOS35866 | CH22_3313FG_646_6_LINK_DJ246D7                          |                 | GENSCAN.1-5   | 1.1 |
|    | 334914 | EOS34845 | CH22_2233FG_457_3_LINK_EM:AC005500                      |                 | GENSCAN.346-2   | 1.1 |
|    | 309527 | EOS09458 | AW150648  | Hs.75621        | protease inhibitor 1 (anti-elastase); alpha-1-antitrypsin   | 1.1 |
| 55 | 318901 | EOS18832 | AW368520  | Hs.24539        | ESTs  | 1.1 |
|    | 320484 | EOS20415 | AA094436  | Hs.155712       | folistatin-like 1   | 1.1 |
|    | 333665 | EOS33596 | CH22_926FG_244_1_LINK_EM:AC005500                       |                 | GENSCAN.99-1  | 1.1 |
|    | 335860 | EOS35791 | CH22_3235FG_629_5_LINK_EM:AC005500                      |                 | GENSCAN.519-4   | 1.1 |
|    | 313339 | EOS13270 | AI682536  | Hs.163495       | ESTs  | 1.1 |
| 60 | 300149 | EOS00080 | AW448916  | Hs.149018       | ESTs  | 1.1 |
|    | 318112 | EOS18043 | AI028162  | Hs.132307       | ESTs  | 1.1 |
|    | 337807 | EOS37738 | CH22_6178FG_LINK_EM:AC005500                            |                 | GENSCAN.9-4   | 1.1 |
|    | 336917 | EOS36848 | CH22_4688FG_346_4_                                      |                 | CH22_FGENES.346-4   | 1.1 |
|    | 337489 | EOS37420 | CH22_5722FG_799_2_                                      |                 | CH22_FGENES.799-2   | 1.1 |
| 65 | 320112 | EOS20043 | T92107  | Hs.188489       | ESTs  | 1.1 |
|    | 332975 | EOS32906 | CH22_199FG_51_10_LINK_EM:AC000097                       |                 | GENSCAN.4-12  | 1.1 |
|    | 327805 | EOS27736 | c_5_hs_gij5867968[ref] gn 2 + 19952 20019 ex 1 2 CDSf   | 9.47 68 988     |   | 1.1 |
|    | 339215 | EOS39146 | CH22_8153FG_LINK_FF113D11                               |                 | GENSCAN.6-10  | 1.1 |
|    | 311965 | EOS11896 | T69279  |                 | EST cluster (not in UniGene)  | 1.1 |
| 75 | 314043 | EOS13974 | AA827082  |                 | EST cluster (not in UniGene)  | 1.1 |
|    | 333447 | EOS33378 | CH22_697FG_154_5_LINK_EM:AC005500                       |                 | GENSCAN.35-6  | 1.1 |
|    | 333242 | EOS33173 | CH22_481FG_111_6_LINK_EM:AC000097                       |                 | GENSCAN.120-5   | 1.1 |
|    | 338596 | EOS38527 | CH22_7343FG_LINK_EM:AC005500                            |                 | GENSCAN.437-2   | 1.1 |
|    | 329989 | EOS29920 | c16_p2_gij4567166[gb]A gn 2 + 72861 73052 ex 1 3 CDSf   | 18.02 192 590   |   | 1.1 |
| 80 | 315675 | EOS15606 | AA652272  | Hs.197320       | ESTs  | 1.1 |
|    | 336722 | EOS36653 | CH22_4245FG_84_2_                                       |                 | CH22_FGENES.84-2  | 1.1 |

|     |        |          |  |     |
|-----|--------|----------|--|-----|
| 5   | 334220 | EOS34151 | CH22_1503FG_359_4_LINK_EM:AC005500.GENSCAN.217-7   |     |
|     |        |          | CH22_FGENES.359_4  | 1.1 |
|     | 336703 | EOS36634 | CH22_4201FG_56_3   | 1.1 |
| 10  | 336397 | EOS36328 | CH22_3812FG_823_12_LINK_BA232E17.GENSCAN.6-11  |     |
|     |        |          | CH22_FGENES.823_12   | 1.1 |
|     | 316105 | EOS16036 | AW295687 Hs.254420 ESTs  | 1.1 |
| 15  | 334661 | EOS34592 | CH22_1969FG_418_9_LINK_EM:AC005500.GENSCAN.279-13  |     |
|     |        |          | CH22_FGENES.418_9  | 1.1 |
|     | 307783 | EOS07714 | AI347274 EST singleton (not in UniGene) with exon hit  | 1.1 |
| 20  | 333997 | EOS33928 | CH22_1275FG_310_22_LINK_EM:AC005500.GENSCAN.167-21   |     |
|     |        |          | CH22_FGENES.310_22   | 1.1 |
|     | 331903 | EOS31834 | AA436673 Hs.29417 Homo sapiens mRNA; cDNA DKFZp586B0323 (from clone DKFZp586B0323)             | 1.1 |
| 25  | 328249 | EOS28180 | c_6_hs gjl6381891[ref] gn 2 - 96352 96527 ex 2 3 CDSi 6.19 176 4550                            |     |
|     |        |          | CH.06_hs gjl6381891  | 1.1 |
|     | 338251 | EOS38182 | CH22_6849FG_LINK_EM:AC005500.GENSCAN.270-1   | 1.1 |
| 30  |        |          | CH22_EM:AC005500.GENSCAN.270-1   | 1.1 |
|     | 323561 | EOS23492 | AA825426 Hs.238832 ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens] | 1.1 |
|     | 301464 | EOS01395 | AA991519 Hs.253324 ESTs  | 1.1 |
| 35  | 335916 | EOS35847 | CH22_3293FG_636_12_LINK_EM:AC005500.GENSCAN.526-12   |     |
|     |        |          | CH22_FGENES.636_12   | 1.1 |
|     | 321828 | EOS21759 | X56197 EST cluster (not in UniGene)  | 1.1 |
| 40  | 327413 | EOS27344 | c_2_hs gjl5867750[ref] gn 3 + 101410 101508 ex 4 5 CDSi 4.34 99 587                            |     |
|     |        |          | CH.02_hs gjl5867750  | 1.1 |
|     | 334474 | EOS34405 | CH22_1773FG_394_5_LINK_EM:AC005500.GENSCAN.257-5   | 1.1 |
| 45  |        |          | CH22_FGENES.394_5  | 1.1 |
|     | 336739 | EOS36670 | CH22_4291FG_117_3  | 1.1 |
|     | 316517 | EOS16448 | AI784315 Hs.123163 ESTs  | 1.1 |
| 50  | 325519 | EOS25450 | c12_hs gjl6017036[ref] gn 5 - 186804 186915 ex 1 3 CDSi 8.36 112 2508                          |     |
|     |        |          | CH.12_hs gjl6017036  | 1.1 |
|     | 333875 | EOS33806 | CH22_1145FG_291_11_LINK_EM:AC005500.GENSCAN.149-6  | 1.1 |
| 55  |        |          | CH22_FGENES.291_11   | 1.1 |
|     | 338221 | EOS38152 | CH22_6797FG_LINK_EM:AC005500.GENSCAN.246-10  | 1.1 |
|     |        |          | CH22_EM:AC005500.GENSCAN.246-10  | 1.1 |
| 60  | 336878 | EOS36809 | CH22_4617FG_318_5  | 1.1 |
|     | 337919 | EOS37850 | CH22_6338FG_LINK_EM:AC005500.GENSCAN.66-5  | 1.1 |
|     |        |          | CH22_EM:AC005500.GENSCAN.66-5  | 1.1 |
| 65  | 309828 | EOS09759 | AW293999 EST singleton (not in UniGene) with exon hit  | 1.1 |
|     | 305259 | EOS05190 | AA679225 EST singleton (not in UniGene) with exon hit  | 1.1 |
|     | 333922 | EOS33853 | CH22_1195FG_296_13_LINK_EM:AC005500.GENSCAN.155-16   | 1.1 |
| 70  |        |          | CH22_FGENES.296_13   | 1.1 |
|     | 322092 | EOS22023 | AF085833 EST cluster (not in UniGene)  | 1.1 |
|     | 313356 | EOS13287 | AI266254 Hs.132929 ESTs  | 1.1 |
| 75  | 318847 | EOS18778 | Z42908 Hs.12308 ESTs   | 1.1 |
|     | 337175 | EOS37106 | CH22_5195FG_567_1  | 1.1 |
|     | 336979 | EOS36910 | CH22_4802FG_385_4  | 1.1 |
| 80  | 312169 | EOS12100 | AI064824 Hs.193385 ESTs  | 1.1 |
|     | 336198 | EOS36129 | CH22_3595FG_719_2_LINK_DA59H18.GENSCAN.21-2  | 1.1 |
|     |        |          | CH22_FGENES.719_2  | 1.1 |
| 85  | 321948 | EOS21879 | AA309612 Hs.118797 ubiquitin-conjugating enzyme E2D 3 (homologous to yeast UBC4/5)             | 1.1 |
|     | 324692 | EOS24623 | AA557952 EST cluster (not in UniGene)  | 1.1 |
|     | 330395 | EOS30326 | D10923 Hs.137555 putative chemokine receptor; GTP-binding protein                              | 1.1 |
| 90  | 333119 | EOS33050 | CH22_347FG_80_4_LINK_EM:AC000097.GENSCAN.65-4  |     |
|     |        |          | CH22_FGENES.80_4   | 1.1 |
|     | 316012 | EOS15943 | AA764950 Hs.119898 ESTs  | 1.1 |
| 95  | 300142 | EOS00073 | AI743419 Hs.205707 ESTs  | 1.1 |
|     | 317215 | EOS17146 | AW014242 Hs.159998 ESTs  | 1.1 |
|     | 329526 | EOS29457 | c10_p2 gjl3983506[gbl]U gn 2 + 12251 12325 ex 3 3 CDSi 7.37 75 178                             |     |
| 100 |        |          | CH.10_p2 gjl3983506  | 1.1 |
|     | 317409 | EOS17340 | AA764968 Hs.4864 KIAA0892 protein  | 1.1 |
|     | 339230 | EOS39161 | CH22_8171FG_LINK_BA354I12.GENSCAN.1-6  | 1.1 |
| 105 |        |          | CH22_BA354I12.GENSCAN.1-6  | 1.1 |
|     | 311598 | EOS11529 | AW023595 Hs.232048 ESTs  | 1.1 |
|     | 339164 | EOS39095 | CH22_8091FG_LINK_DA59H18.GENSCAN.69-4  | 1.1 |
| 110 |        |          | CH22_DA59H18.GENSCAN.69-4  | 1.1 |
|     | 326725 | EOS26656 | c20_hs gjl6552456[ref] gn 2 - 223005 223125 ex 5 6 CDSi 6.10 121 1038                          |     |
|     |        |          | CH.20_hs gjl6552456  | 1.1 |
| 115 | 330952 | EOS30883 | H02855 Hs.29567 ESTs   | 1.1 |
|     | 334621 | EOS34552 | CH22_1928FG_412_4_LINK_EM:AC005500.GENSCAN.275-4   |     |
|     |        |          | CH22_FGENES.412_4  | 1.1 |
| 120 | 301685 | EOS01616 | W67730 EST cluster (not in UniGene) with exon hit  | 1.1 |
|     | 308781 | EOS08712 | AI811707 EST singleton (not in UniGene) with exon hit  | 1.1 |
|     | 323413 | EOS23344 | AA248828 Hs.225676 ESTs  | 1.1 |
| 125 | 306723 | EOS06654 | AI026151 EST singleton (not in UniGene) with exon hit  | 1.1 |
|     | 331258 | EOS31189 | Z41777 Hs.27413 ESTs   | 1.1 |
|     | 313028 | EOS12959 | AI355433 Hs.190856 ESTs  | 1.1 |
| 130 | 333002 | EOS32933 | CH22_226FG_59_3_LINK_EM:AC000097.GENSCAN.21-3  |     |
|     |        |          | CH22_FGENES.59_3   | 1.1 |
|     | 303011 | EOS02942 | AF090405 EST cluster (not in UniGene) with exon hit  | 1.1 |
| 135 | 317687 | EOS17618 | AA972990 Hs.127904 ESTs  | 1.1 |
|     | 328779 | EOS28710 | c_7_hs gjl5868309[ref] gn 4 + 41570 41639 ex 1 5 CDSf 2.65 70 5365                             |     |
|     |        |          | CH.07_hs gjl5868309  | 1.1 |
| 140 | 338707 | EOS38638 | CH22_7487FG_LINK_EM:AC005500.GENSCAN.482-2   |     |
|     |        |          | CH22_EM:AC005500.GENSCAN.482-2   | 1.1 |
|     | 337974 | EOS37905 | CH22_6427FG_LINK_EM:AC005500.GENSCAN.106-3   | 1.1 |
| 145 |        |          | CH22_EM:AC005500.GENSCAN.106-3   | 1.1 |
|     | 332854 | EOS32785 | CH22_71FG_22_1_LINK_C20H12.GENSCAN.15-2  |     |

|    |        |          |  |  |     |
|----|--------|----------|--|--|-----|
|    |        |          |  | CH22_FGENES.22_1   | 1.1 |
|    | 311225 | EOS11156 | AW451982   | Hs.248613 ESTs   | 1.1 |
|    | 337094 | EOS37025 | CH22_5018FG_465_19_                                | CH22_FGENES.465-19   | 1.1 |
|    | 319357 | EOS19288 | F13425   | Hs.26229 ESTs  | 1.1 |
| 5  | 332958 | EOS32889 | CH22_182FG_48_15_LINK_EM:AC000097.GENSCAN.2-11     |  | 1.1 |
|    |        |          |  | CH22_FGENES.48_15  | 1.1 |
|    | 309634 | EOS09565 | AW193825   | EST singleton (not in UniGene) with exon hit   | 1.1 |
|    | 321171 | EOS21102 | AI769410   | Hs.221461 ESTs   | 1.1 |
| 10 | 316440 | EOS16371 | AI954795   | Hs.156135 ESTs   | 1.1 |
|    | 311665 | EOS11596 | AW294254   | Hs.223742 ESTs   | 1.1 |
|    | 327548 | EOS27479 | c_3_hs gij5867797[ref] gn 2                        | - 81067 81130 ex 3 7 CDSi 6.42 64 12   | 1.1 |
|    |        |          |  | CH.03_hs gij5867797  | 1.1 |
|    | 314940 | EOS14871 | AW452768   | Hs.162045 ESTs   | 1.1 |
| 15 | 326401 | EOS26332 | c19_hs gij5867355[ref] gn 1                        | + 35165 35332 ex 9 11 CDSi 0.41 168 788  | 1.1 |
|    |        |          |  | CH.19_hs gij5867355  | 1.1 |
|    | 336347 | EOS36278 | CH22_3759FG_815_3_LINK_BA232E17.GENSCAN.1-24       |  | 1.1 |
|    |        |          |  | CH22_FGENES.815_3  | 1.1 |
|    | 322297 | EOS22228 | W76548   | Hs.136026 ESTs; Moderately similar to !!!! ALU SUBFAMILY SC WARNING ENTRY !!!! [H.sapiens]         | 1.1 |
| 20 | 309977 | EOS09908 | AW451663   | EST singleton (not in UniGene) with exon hit   | 1.1 |
|    | 333466 | EOS33397 | CH22_717FG_161_2_LINK_EM:AC005500.GENSCAN.42-2     |  | 1.1 |
|    |        |          |  | CH22_FGENES.161_2  | 1.1 |
|    | 329170 | EOS29101 | c_x_hs gij5868693[ref] gn 2                        | + 67924 68019 ex 6 8 CDSi 3.30 96 1882   | 1.1 |
|    |        |          |  | CH.X_hs gij5868693   | 1.1 |
| 25 | 329479 | EOS29410 | c10_p2 gij3983526[gb]A gn 3                        | - 7425 7561 ex 1 3 CDSi 4.33 137 22  | 1.1 |
|    |        |          |  | CH.10_p2 gij3983526  | 1.1 |
|    | 326668 | EOS26599 | c20_hs gij6552455[ref] gn 1                        | + 146726 146838 ex 11 11 CDSi 1.84 113 767   | 1.1 |
|    |        |          |  | CH.20_hs gij6552455  | 1.1 |
|    | 319364 | EOS19295 | H06538   | Hs.12270 ESTs  | 1.1 |
| 30 | 302988 | EOS02919 | W23986   | Hs.34578 alpha2;3-sialyltransferase  | 1.1 |
|    | 327687 | EOS27618 | c_4_hs gij5867847[ref] gn 1                        | - 169293 169362 ex 2 3 CDSi -0.28 70 782   | 1.1 |
|    |        |          |  | CH.04_hs gij5867847  | 1.1 |
|    | 339413 | EOS39344 | CH22_8405FG_LINK_DJ579N16.GENSCAN.5-8              |  | 1.1 |
|    |        |          |  | CH22_DJ579N16.GENSCAN.5-8  | 1.1 |
| 35 | 306156 | EOS06087 | AA918274   | Hs.76067 heat shock 27kD protein 1   | 1.1 |
|    | 320858 | EOS20789 | D59968   | EST cluster (not in UniGene)   | 1.1 |
|    | 325447 | EOS25378 | c12_hs gij5866941[ref] gn 3                        | - 372480 372621 ex 2 3 CDSi 9.16 142 1026  | 1.1 |
|    |        |          |  | CH.12_hs gij5866941  | 1.1 |
|    | 322696 | EOS22627 | AI064724   | Hs.228468 ESTs   | 1.1 |
| 40 | 329959 | EOS29890 | c16_p2 gij5103803[gb]A gn 3                        | + 188050 188193 ex 8 8 CDSi 2.01 144 361   | 1.1 |
|    |        |          |  | CH.16_p2 gij5103803  | 1.1 |
|    | 312628 | EOS12559 | AA632817   | Hs.190316 ESTs   | 1.1 |
|    | 339305 | EOS39236 | CH22_8262FG_LINK_BA354I12.GENSCAN.21-3             |  | 1.1 |
|    |        |          |  | CH22_BA354I12.GENSCAN.21-3   | 1.1 |
| 45 | 311829 | EOS11760 | AI078483   | Hs.134549 ESTs   | 1.1 |
|    | 303270 | EOS03201 | AL120518   | Hs.105352 ESTs   | 1.1 |
|    | 321226 | EOS21157 | AA311443   | Hs.251416 Homo sapiens mRNA; cDNA DKFZp586E2317 (from clone DKFZp586E2317)                         | 1.1 |
|    | 335827 | EOS35758 | CH22_3200FG_620_1_LINK_EM:AC005500.GENSCAN.512-1   |  | 1.1 |
|    |        |          |  | CH22_FGENES.620_1  | 1.1 |
| 50 | 336677 | EOS36608 | CH22_4155FG_43_5_                                  | CH22_FGENES.43-5   | 1.1 |
|    | 330081 | EOS30012 | c19_p2 gij6015314[gb]A gn 1                        | - 5768 5835 ex 4 9 CDSi 2.88 68 162  | 1.1 |
|    |        |          |  | CH.19_p2 gij6015314  | 1.1 |
|    | 339313 | EOS39244 | CH22_8272FG_LINK_BA354I12.GENSCAN.22-11            |  | 1.1 |
|    |        |          |  | CH22_BA354I12.GENSCAN.22-11  | 1.1 |
| 55 | 319936 | EOS19867 | W22152   | EST cluster (not in UniGene)   | 1.1 |
|    | 332858 | EOS32789 | CH22_76FG_24_1_LINK_C20H12.GENSCAN.16-6            |  | 1.1 |
|    |        |          |  | CH22_FGENES.24_1   | 1.1 |
|    | 315630 | EOS15561 | AA648355   | Hs.185155 ESTs; Weakly similar to echinoderm microtubule-associated protein-like EMAP2 [H.sapiens] | 1.1 |
|    | 332995 | EOS32926 | CH22_219FG_58_2_LINK_EM:AC000097.GENSCAN.19-2      |  | 1.1 |
| 60 |        |          |  | CH22_FGENES.58_2   | 1.1 |
|    | 333441 | EOS33372 | CH22_691FG_151_5_LINK_EM:AC005500.GENSCAN.32-5     |  | 1.1 |
|    |        |          |  | CH22_FGENES.151_5  | 1.1 |
|    | 333496 | EOS33427 | CH22_748FG_168_6_LINK_EM:AC005500.GENSCAN.47-5     |  | 1.1 |
|    |        |          |  | CH22_FGENES.168_6  | 1.1 |
| 65 | 339188 | EOS39119 | CH22_8123FG_LINK_DA59H18.GENSCAN.72-16             |  | 1.1 |
|    |        |          |  | CH22_DA59H18.GENSCAN.72-16   | 1.1 |
|    | 336981 | EOS36912 | CH22_4818FG_397_7_                                 | CH22_FGENES.397-7  | 1.1 |
|    | 312142 | EOS12073 | AW298359   | Hs.221069 ESTs   | 1.1 |
|    | 315779 | EOS15710 | AW015736   | Hs.211378 ESTs   | 1.1 |
| 70 | 318596 | EOS18527 | AI470235   | Hs.172698 EST  | 1.1 |
|    | 335701 | EOS35632 | CH22_3062FG_599_1_LINK_EM:AC005500.GENSCAN.490-2   |  | 1.1 |
|    |        |          |  | CH22_FGENES.599_1  | 1.1 |
|    | 319395 | EOS19326 | AW062570   | Hs.13809 ESTs  | 1.1 |
|    | 304236 | EOS04167 | W93278   | EST singleton (not in UniGene) with exon hit   | 1.1 |
| 75 | 307264 | EOS07195 | AI202211   | EST singleton (not in UniGene) with exon hit   | 1.1 |
|    | 334066 | EOS33997 | CH22_1344FG_327_21_LINK_EM:AC005500.GENSCAN.181-23 |  | 1.1 |
|    |        |          |  | CH22_FGENES.327_21   | 1.1 |
|    | 327042 | EOS26973 | c21_hs gij6531965[ref] gn 18                       | - 1380806 1381443 ex 1 5 CDSi 30.85 638 943  | 1.1 |
|    |        |          |  | CH.21_hs gij6531965  | 1.1 |
| 80 | 326025 | EOS25956 | c17_hs gij5867176[ref] gn 1                        | + 70854 70915 ex 6 8 CDSi -1.46 62 127   | 1.1 |
|    |        |          |  | CH.17_hs gij5867176  | 1.1 |
|    | 325609 | EOS25540 | c14_hs gij5866996[ref] gn 28                       | - 981751 981849 ex 1 10 CDSi 1.46 99 101   | 1.1 |
|    |        |          |  | CH.14_hs gij5866996  | 1.1 |
|    | 319983 | EOS19914 | T81429   | EST cluster (not in UniGene)   | 1.1 |
| 85 | 334298 | EOS34229 | CH22_1589FG_372_4_LINK_EM:AC005500.GENSCAN.232-5   |  | 1.1 |
|    |        |          |  | CH22_FGENES.372_4  | 1.1 |
|    | 323203 | EOS23134 | AA203135   | Hs.130186 ESTs   | 1.1 |

|    |        |          |  |   |     |
|----|--------|----------|--|---|-----|
|    | 305700 | EOS05631 | AA815428   | EST singleton (not in UniGene) with exon hit                                | 1.1 |
|    | 313304 | EOS13235 | AI334078 Hs.152438   | ESTs  | 1.1 |
|    | 310716 | EOS10647 | AI589618 Hs.192413   | ESTs  | 1.1 |
| 5  | 327049 | EOS26980 | c21_hs gij 6531965 ref  gn 24 - 1924026 1924110 ex 2 6 CDSi 9.43 85 1012 | CH.21_hs gij 6531965  | 1.1 |
|    | 313749 | EOS13680 | AW450376 Hs.130803   | ESTs  | 1.1 |
|    | 307041 | EOS06972 | AI144243   | EST singleton (not in UniGene) with exon hit                                | 1.1 |
|    | 322394 | EOS22325 | AF077208   | EST cluster (not in UniGene)  | 1.1 |
| 10 | 326416 | EOS26347 | c19_hs gij 5867362 ref  gn 3 - 45283 45375 ex 3 3 CDSf 5.65 93 923       | CH.19_hs gij 5867362  | 1.1 |
|    | 333947 | EOS33878 | CH22_1221FG_303_1_LINK_EM:AC005500.GENSCAN.162-5                         | CH22_FGENES.303_1   | 1.1 |
|    | 324609 | EOS24540 | AW299534   | EST cluster (not in UniGene)  | 1.1 |
| 15 | 330057 | EOS29988 | c17_p2 gij 6478962 gb A gn 3 + 75145 75287 ex 3 3 CDSi -2.56 143 150     | CH.17_p2 gij 6478962  | 1.1 |
|    | 337603 | EOS37534 | CH22_5896FG_LINK_C20H12.GENSCAN.16-2                                     | CH22_C20H12.GENSCAN.16-2  | 1.1 |
|    | 332913 | EOS32844 | CH22_134FG_36_18_LINK_C20H12.GENSCAN.28-17                               | CH22_FGENES.36_18   | 1.1 |
| 20 | 310026 | EOS09957 | T24895 Hs.100691   | ESTs  | 1.1 |
|    | 330153 | EOS30084 | c21_p2 gij 4325335 gb A gn 2 + 146951 147475 ex 2 2 CDSi 25.45 525 233   | CH.21_p2 gij 4325335  | 1.1 |
|    | 334118 | EOS34049 | CH22_1396FG_330_19_LINK_EM:AC005500.GENSCAN.185-20                       | CH22_FGENES.330_19  | 1.1 |
| 25 | 324795 | EOS24726 | AI494481 Hs.141579   | ESTs  | 1.1 |
|    | 332530 | EOS32461 | M31682 Hs.1735   | inhibin; beta B (activin AB beta polypeptide)                               | 1.1 |
|    | 332048 | EOS31979 | AA496019 Hs.201591   | ESTs  | 1.1 |
|    | 334532 | EOS34463 | CH22_1834FG_402_13_LINK_EM:AC005500.GENSCAN.266-13                       | CH22_FGENES.402_13  | 1.1 |
| 30 | 329762 | EOS29693 | c14_p2 gij 6048280 emb  gn 3 + 127744 127878 ex 2 4 CDSi 11.66 135 1054  | CH.14_p2 gij 6048280  | 1.1 |
|    | 332909 | EOS32840 | CH22_130FG_36_13_LINK_C20H12.GENSCAN.28-10                               | CH22_FGENES.36_13   | 1.1 |
| 35 | 321253 | EOS21184 | AI699484   | EST cluster (not in UniGene)  | 1.1 |
|    | 336572 | EOS36503 | CH22_4007FG_843_12_LINK_DJ579N16.GENSCAN.15-13                           | CH22_FGENES.843_12  | 1.1 |
|    | 328768 | EOS28699 | c_7_hs gij 6017031 ref  gn 5 - 223741 224238 ex 1 1 CDSO 30.00 498 5285  | CH.07_hs gij 6017031  | 1.1 |
| 40 | 334335 | EOS34266 | CH22_1627FG_375_12_LINK_EM:AC005500.GENSCAN.235-12                       | CH22_FGENES.375_12  | 1.1 |
|    | 334063 | EOS33994 | CH22_1341FG_327_17_LINK_EM:AC005500.GENSCAN.181-20                       | CH22_FGENES.327_17  | 1.1 |
|    | 333011 | EOS32942 | CH22_235FG_61_3_LINK_EM:AC000097.GENSCAN.23-3                            | CH22_FGENES.61_3  | 1.1 |
| 45 | 304677 | EOS04608 | AA548071   | EST singleton (not in UniGene) with exon hit                                | 1.1 |
|    | 313948 | EOS13879 | AW452823 Hs.135268   | ESTs  | 1.1 |
|    | 334358 | EOS34289 | CH22_1652FG_378_1_LINK_EM:AC005500.GENSCAN.239-1                         | CH22_FGENES.378_1   | 1.1 |
| 50 | 328479 | EOS28410 | c_7_hs gij 5868449 ref  gn 1 - 331 560 ex 1 31 CDSi 18.51 230 2100       | CH.07_hs gij 5868449  | 1.1 |
|    | 335813 | EOS35744 | CH22_3185FG_618_1_LINK_EM:AC005500.GENSCAN.510-1                         | CH22_FGENES.618_1   | 1.1 |
|    | 312430 | EOS12361 | AW139117 Hs.117494   | ESTs  | 1.1 |
| 55 | 324783 | EOS24714 | AA640770   | EST cluster (not in UniGene)  | 1.1 |
|    | 337776 | EOS37707 | CH22_6132FG_LINK_EM:AC000097.GENSCAN.119-18                              | CH22_EM:AC000097.GENSCAN.119-18   | 1.1 |
|    | 327205 | EOS27136 | c_1_hs gij 5867447 ref  gn 5 + 167335 167576 ex 9 9 CDSi 15.50 242 259   | CH.01_hs gij 5867447  | 1.1 |
| 60 | 315198 | EOS15129 | AI741506 Hs.186753   | ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens] | 1.1 |
|    | 336135 | EOS36066 | CH22_3525FG_704_3_LINK_DA59H18.GENSCAN.9-5                               | CH22_FGENES.704_3   | 1.1 |
|    | 318558 | EOS18489 | AW402677 Hs.90372  | ESTs  | 1.1 |
|    | 328152 | EOS28083 | c_6_hs gij 5868060 ref  gn 1 - 73981 74203 ex 1 8 CDSi 31.69 223 3411    | CH.06_hs gij 5868060  | 1.1 |
| 65 | 330211 | EOS30142 | c_5_p2 gij 6013592 gb A gn 1 + 59158 59215 ex 2 4 CDSi 4.20 58 184       | CH.05_p2 gij 6013592  | 1.1 |
|    | 339280 | EOS39211 | CH22_8234FG_LINK_BA354I12.GENSCAN.14-12                                  | CH22_BA354I12.GENSCAN.14-12   | 1.1 |
| 70 | 332045 | EOS1976  | AA491253 Hs.155045   | bromodomain adjacent to zinc finger domain; 2A                              | 1.1 |
|    | 313597 | EOS13528 | AW162263 Hs.249990   | ESTs  | 1.1 |
|    | 329503 | EOS29434 | c10_p2 gij 3983517 gb U gn 2 - 1801 1937 ex 1 4 CDSi 4.33 137 101        | CH.10_p2 gij 3983517  | 1.1 |
|    | 333488 | EOS33419 | CH22_740FG_167_3_LINK_EM:AC005500.GENSCAN.46-10                          | CH22_FGENES.167_3   | 1.1 |
| 75 | 311960 | EOS11891 | AW440133 Hs.189690   | ESTs  | 1.1 |
|    | 320590 | EOS20521 | U67058 Hs.168102   | Human proteinase activated receptor-2 mRNA; 3'UTR                           | 1.1 |
|    | 334047 | EOS33978 | CH22_1325FG_326_5_LINK_EM:AC005500.GENSCAN.175-5                         | CH22_FGENES.326_5   | 1.1 |
| 80 | 304782 | EOS04713 | AA582081   | EST singleton (not in UniGene) with exon hit                                | 1.1 |
|    | 324231 | EOS24162 | W60827   | EST cluster (not in UniGene)  | 1.1 |
|    | 327212 | EOS27143 | c_1_hs gij 5867463 ref  gn 1 - 42308 42424 ex 5 13 CDSi 6.58 117 325     | CH.01_hs gij 5867463  | 1.1 |
|    | 335857 | EOS35788 | CH22_3232FG_629_1_LINK_EM:AC005500.GENSCAN.519-1                         | CH22_FGENES.629_1   | 1.1 |
| 85 | 317775 | EOS17706 | AA974603 Hs.181123   | ESTs  | 1.1 |
|    | 331053 | EOS30984 | N70242 Hs.183146   | ESTs  | 1.1 |

|    |        |          |   |   |     |
|----|--------|----------|---|---|-----|
|    | 335940 | EOS35871 | CH22_3318FG_646_13_LINK_DJ246D7.GENSCAN.1-12                            |   |     |
|    |        |          | CH22_FGENES.646_13  | 1.1   |     |
|    | 322568 | EOS22499 | W87342 Hs.209652  | ESTs  | 1.1 |
|    | 314091 | EOS14022 | AI253112 Hs.133540  | ESTs  | 1.1 |
| 5  | 313570 | EOS13501 | AA041455 Hs.209312  | ESTs  | 1.1 |
|    | 300967 | EOS00898 | AA565209 Hs.190216  | ESTs  | 1.1 |
|    | 314544 | EOS14475 | AA399018 Hs.250835  | ESTs  | 1.1 |
|    | 328321 | EOS28252 | c_7_hs_gij 5868373 ref  gn 7 - 1029614 1029673 ex 1 3 CDSI -2.40 60 448 |   |     |
|    |        |          | CH.07_hs_gij 5868373  | 1.1   |     |
| 10 | 310979 | EOS10910 | AW445166 Hs.170802  | ESTs  | 1.1 |
|    | 310730 | EOS10661 | AI939421 Hs.160900  | ESTs  | 1.1 |
|    | 318471 | EOS18402 | AW137725 Hs.146874  | ESTs  | 1.1 |
|    | 315533 | EOS15464 | AW206191 Hs.152774  | ESTs  | 1.1 |
| 15 | 325751 | EOS25682 | c14_hs_gij 6682474 ref  gn 4 + 130437 130520 ex 6 7 CDSI 0.22 84 1666   |   |     |
|    |        |          | CH.14_hs_gij 6682474  | 1.1   |     |
|    | 318780 | EOS18711 | R90906 Hs.113307  | ESTs  | 1.1 |
|    | 313271 | EOS13202 | AW444819 Hs.144851  | ESTs; Weakly similar to C09F5.2 [C.elegans]                                       | 1.1 |
|    | 304546 | EOS04477 | AA486074  | EST singleton (not in UniGene) with exon hit                                      | 1.1 |
| 20 | 330618 | EOS30549 | X55990 Hs.73839   | ribonuclease; RNase A family; 3 (eosinophil cationic protein)                     | 1.1 |
|    | 332931 | EOS32862 | CH22_152FG_38_5_LINK_C20H12.GENSCAN.29-5                                |   |     |
|    |        |          | CH22_FGENES.38_5  | 1.1   |     |
|    | 336602 | EOS36533 | CH22_4047FG_372_4_LINK_EM:AC005500.GENSCAN.232-4                        |   |     |
|    |        |          | CH22_FGENES.372_4   | 1.1   |     |
| 25 | 311185 | EOS11116 | AI638294 Hs.224665  | ESTs  | 1.1 |
|    | 337585 | EOS37516 | CH22_5873FG_LINK_C20H12.GENSCAN.5-3                                     |   |     |
|    |        |          | CH22_C20H12.GENSCAN.5-3   | 1.1   |     |
|    | 310249 | EOS10180 | AW071751 Hs.13179   | ESTs; Moderately similar to !!!! ALU SUBFAMILY SQ WARNING ENTRY !!!! [H.sapiens]  | 1.1 |
|    | 314578 | EOS14509 | AA410183 Hs.137475  | ESTs  | 1.1 |
| 30 | 310750 | EOS10681 | AI373163 Hs.170333  | ESTs  | 1.1 |
|    | 333968 | EOS33899 | CH22_1245FG_307_4_LINK_EM:AC005500.GENSCAN.165-5                        |   |     |
|    |        |          | CH22_FGENES.307_4   | 1.1   |     |
|    | 316133 | EOS16064 | AI187742 Hs.125562  | ESTs  | 1.1 |
|    | 308337 | EOS08268 | AI608947  | EST singleton (not in UniGene) with exon hit                                      | 1.1 |
| 35 | 326160 | EOS26091 | c17_hs_gij 5867254 ref  gn 6 - 112000 112137 ex 2 4 CDSi 8.01 138 1952  |   |     |
|    |        |          | CH.17_hs_gij 5867254  | 1.1   |     |
|    | 336023 | EOS35954 | CH22_3406FG_669_12_LINK_DJ32110.GENSCAN.9-17                            |   |     |
|    |        |          | CH22_FGENES.669_12  | 1.1   |     |
|    | 323479 | EOS23410 | AA278246  | EST cluster (not in UniGene)  | 1.1 |
| 40 | 336090 | EOS36021 | CH22_3477FG_689_2_LINK_DJ32110.GENSCAN.23-20                            |   |     |
|    |        |          | CH22_FGENES.689_2   | 1.1   |     |
|    | 311192 | EOS11123 | AW237220 Hs.211130  | ESTs  | 1.1 |
|    | 335081 | EOS35012 | CH22_2409FG_488_4_LINK_EM:AC005500.GENSCAN.384-6                        |   |     |
|    |        |          | CH22_FGENES.488_4   | 1.1   |     |
| 45 | 309519 | EOS09450 | AW148940 Hs.248647  | EST   | 1.1 |
|    | 321172 | EOS21103 | H49160 Hs.133472  | ESTs  | 1.1 |
|    | 301976 | EOS01907 | T97905  | EST cluster (not in UniGene) with exon hit  | 1.1 |
|    | 323012 | EOS22943 | AI832201 Hs.211469  | ESTs  | 1.1 |
|    | 319528 | EOS19459 | R08673 Hs.177514  | ESTs  | 1.1 |
| 50 | 329838 | EOS29769 | c14_p2_gij 6672062 emb  gn 2 + 33990 34098 ex 3 4 CDSi 9.11 109 2222    |   |     |
|    |        |          | CH.14_p2_gij 6672062  | 1.1   |     |
|    | 302623 | EOS02554 | AB019571  | EST cluster (not in UniGene) with exon hit  | 1.1 |
|    | 334433 | EOS34364 | CH22_1731FG_385_8_LINK_EM:AC005500.GENSCAN.249-6                        |   |     |
|    |        |          | CH22_FGENES.385_8   | 1.1   |     |
| 55 | 304747 | EOS04678 | AA577816  | EST singleton (not in UniGene) with exon hit                                      | 1.1 |
|    | 333270 | EOS33201 | CH22_513FG_121_1_LINK_EM:AC005500.GENSCAN.4-11                          |   |     |
|    |        |          | CH22_FGENES.121_1   | 1.1   |     |
|    | 307054 | EOS06985 | AI148181 Hs.176835  | EST   | 1.1 |
|    | 320764 | EOS20695 | R73070 Hs.246927  | ESTs  | 1.1 |
| 60 | 321523 | EOS21454 | H78472 Hs.191325  | ESTs; Weakly similar to cDNA EST yk414c9.3 comes from this gene [C.elegans]       | 1.1 |
|    | 322114 | EOS22045 | AA643791 Hs.191740  | ESTs  | 1.1 |
|    | 303582 | EOS03513 | AA377444  | EST cluster (not in UniGene) with exon hit  | 1.1 |
|    | 322924 | EOS22855 | AA669253 Hs.193971  | ESTs  | 1.1 |
|    | 311179 | EOS11110 | AI880843 Hs.223333  | ESTs  | 1.1 |
| 65 | 318601 | EOS18532 | T39921  | EST cluster (not in UniGene)  | 1.1 |
|    | 309791 | EOS09722 | AW276176 Hs.73742   | ribosomal protein; large; P0  | 1.1 |
|    | 333882 | EOS33813 | CH22_1153FG_292_4_LINK_EM:AC005500.GENSCAN.150-4                        |   |     |
|    |        |          | CH22_FGENES.292_4   | 1.1   |     |
|    | 337645 | EOS37576 | CH22_5960FG_LINK_EM:AC000097.GENSCAN.10-8                               |   |     |
|    |        |          | CH22_EM:AC000097.GENSCAN.10-8   | 1.1   |     |
| 70 | 335623 | EOS35554 | CH22_2983FG_584_2_LINK_EM:AC005500.GENSCAN.478-2                        |   |     |
|    |        |          | CH22_FGENES.584_2   | 1.1   |     |
|    | 314745 | EOS14676 | AA564489 Hs.137526  | ESTs  | 1.1 |
|    | 330790 | EOS30721 | T48536 Hs.105807  | ESTs  | 1.1 |
| 75 | 332071 | EOS32002 | AA598594 Hs.112475  | ESTs  | 1.1 |
|    | 312005 | EOS11936 | T78460 Hs.13941   | ESTs  | 1.1 |
|    | 330694 | EOS30625 | AA019806 Hs.108447  | spinocerebellar ataxia 7 (olivopontocerebellar atrophy with retinal degeneration) | 1.1 |
|    | 330739 | EOS30670 | AA293477 Hs.227591  | ESTs  | 1.1 |
|    | 303042 | EOS02973 | AF129532  | EST cluster (not in UniGene) with exon hit  | 1.1 |
| 80 | 323091 | EOS23022 | AW014094 Hs.210761  | ESTs  | 1.1 |
|    | 328820 | EOS28751 | c_7_hs_gij 5868330 ref  gn 1 + 90446 90602 ex 3 4 CDSi 10.20 157 5634   |   |     |
|    |        |          | CH.07_hs_gij 5868330  | 1.1   |     |
|    | 300472 | EOS00403 | T90622 Hs.82609   | hydroxymethylbilane synthase  | 1.1 |
|    | 310645 | EOS10576 | AI420742 Hs.163502  | ESTs  | 1.1 |
| 85 | 332238 | EOS32169 | N53480 Hs.108622  | ESTs  | 1.1 |
|    | 300966 | EOS00897 | AA564740 Hs.258401  | ESTs  | 1.1 |
|    | 330437 | EOS30368 | HG2730-HT2827   | Fibrinogen, A Alpha Polypeptide, Alt. Splice 2, E                                 | 1.1 |

|    |        |          |  |  |     |
|----|--------|----------|--|--|-----|
|    | 302292 | EOS02223 | AF067797   | EST cluster (not in UniGene) with exon hit   | 1.1 |
|    | 330138 | EOS30069 | c21_p2 gij4210430[emb] gn 1 - 22334 22460 ex 3 3 CDSf 16.56 127 105      | CH.21_p2 gij4210430  | 1.1 |
| 5  | 332952 | EOS32883 | CH22_176FG_48_8_LINK_EM:AC000097.GENSCAN.2-4                             | CH22_FGENES.48_8   | 1.1 |
|    | 319901 | EOS19832 | T77136   | Hs.8765 RNA helicase-related protein   | 1.1 |
|    | 321166 | EOS21097 | AA411263   | Hs.128783 ESTs   | 1.1 |
|    | 336227 | EOS36158 | CH22_3625FG_730_2_LINK_DA59H18.GENSCAN.36-2                              | CH22_FGENES.730_2  | 1.1 |
| 10 | 302332 | EOS02263 | AI833168   | Hs.184507 Homo sapiens Chromosome 16 BAC clone CIT987SK-A-328A3                            | 1.1 |
|    | 313800 | EOS13731 | AW296132   | Hs.166674 ESTs   | 1.1 |
|    | 339356 | EOS39287 | CH22_8326FG_LINK_BA354112.GENSCAN.31-1                                   | CH22_BA354112.GENSCAN.31-1   | 1.1 |
| 15 | 324512 | EOS24443 | AW502125   | EST cluster (not in UniGene)   | 1.1 |
|    | 319235 | EOS19166 | F11330   | Hs.177633 ESTs   | 1.1 |
|    | 320352 | EOS20283 | Y13323   | Hs.145296 disintegrin protease   | 1.1 |
|    | 338316 | EOS38247 | CH22_6944FG_LINK_EM:AC005500.GENSCAN.304-2                               | CH22_EM:AC005500.GENSCAN.304-2   | 1.1 |
| 20 | 333964 | EOS33895 | CH22_1241FG_305_2_LINK_EM:AC005500.GENSCAN.164-2                         | CH22_FGENES.305_2  | 1.1 |
|    | 312758 | EOS12689 | AA721107   | Hs.202604 ESTs   | 1.1 |
|    | 338178 | EOS38109 | CH22_6726FG_LINK_EM:AC005500.GENSCAN.219-6                               | CH22_EM:AC005500.GENSCAN.219-6   | 1.1 |
| 25 | 315199 | EOS15130 | AA877996   | Hs.125376 ESTs   | 1.1 |
|    | 312321 | EOS12252 | R66210   | Hs.186937 ESTs   | 1.1 |
|    | 338765 | EOS38696 | CH22_7588FG_LINK_EM:AC005500.GENSCAN.518-1                               | CH22_EM:AC005500.GENSCAN.518-1   | 1.1 |
| 30 | 330547 | EOS30478 | U32989   | Hs.183671 tryptophan 2;3-dioxygenase   | 1.1 |
|    | 315368 | EOS15299 | AW291563   | Hs.152495 ESTs   | 1.1 |
|    | 328691 | EOS28622 | c.7_hs gij6588001[ref] gn 7 - 579598 579664 ex 2 3 CDSi 12.78 67 4326    | CH.07_hs gij6588001  | 1.1 |
|    | 329179 | EOS29110 | c.x_hs gij5868704[ref] gn 2 + 181639 181815 ex 3 4 CDSi 0.32 177 1939    | CH.X_hs gij5868704   | 1.1 |
| 35 | 327072 | EOS27003 | c21_hs gij6531965[ref] gn 55 - 3796429 3797197 ex 4 4 CDSf 9.33 769 1270 | CH.21_hs gij6531965  | 1.1 |
|    | 312056 | EOS11987 | T83748   | Hs.189712 ESTs   | 1.1 |
|    | 339128 | EOS39059 | CH22_8046FG_LINK_DA59H18.GENSCAN.55-2                                    | CH22_DA59H18.GENSCAN.55-2  | 1.1 |
| 40 | 307646 | EOS07577 | AI302236   | EST singleton (not in UniGene) with exon hit   | 1.1 |
|    | 319198 | EOS19129 | F07354   | EST cluster (not in UniGene)   | 1.1 |
|    | 338556 | EOS38487 | CH22_7283FG_LINK_EM:AC005500.GENSCAN.417-8                               | CH22_EM:AC005500.GENSCAN.417-8   | 1.1 |
| 45 | 306143 | EOS06074 | AA916314   | EST singleton (not in UniGene) with exon hit   | 1.1 |
|    | 332384 | EOS32315 | M11433   | Hs.101850 retinol-binding protein 1; cellular  | 1.1 |
|    | 325100 | EOS25031 | T10265   | Hs.116122 ESTs; Weakly similar to coded for by C. elegans cDNA yk30b3.5 [C.elegans]        | 1.1 |
|    | 309839 | EOS09770 | AW296076   | EST singleton (not in UniGene) with exon hit   | 1.1 |
|    | 312180 | EOS12111 | AI248285   | Hs.118348 ESTs   | 1.1 |
|    | 330385 | EOS30316 | AA449749   | Hs.31386 ESTs; Highly similar to secreted apoptosis related protein 1 [H.sapiens]          | 1.1 |
| 50 | 315882 | EOS15813 | AI831297   | Hs.123310 ESTs   | 1.1 |
|    | 325843 | EOS25774 | c16_hs gij6552453[ref] gn 1 - 7126 7232 ex 1 3 CDSi 1.87 107 182         | CH.16_hs gij6552453  | 1.1 |
|    | 330783 | EOS30714 | D60050   | Hs.34812 ESTs  | 1.1 |
|    | 317224 | EOS17155 | D56760   | Hs.8122 ESTs   | 1.1 |
| 55 | 316042 | EOS15973 | AW297979   | Hs.170698 ESTs   | 1.1 |
|    | 333524 | EOS33455 | CH22_781FG_175_10_LINK_EM:AC005500.GENSCAN.53-15                         | CH22_FGENES.175_10   | 1.1 |
|    | 302357 | EOS02288 | X03178   | Hs.198246 group-specific component (vitamin D binding protein)                             | 1.1 |
|    | 309830 | EOS09761 | AW294725   | EST singleton (not in UniGene) with exon hit   | 1.1 |
| 60 | 321489 | EOS21420 | AW392474   | Hs.172759 ESTs; Moderately similar to !!!! ALU SUBFAMILY SQ WARNING ENTRY !!!! [H.sapiens] | 1.1 |
|    | 312304 | EOS12235 | AA491949   | Hs.183359 ESTs   | 1.1 |
|    | 322026 | EOS21957 | AA233527   | Hs.213289 low density lipoprotein receptor (familial hypercholesterolemia)                 | 1.1 |

TABLE 1A

Table 1A shows the accession numbers for those primekeys in Table 1 which lack a unigeneID. Listed for each probeset is the gene cluster (CAT) number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number  
 CAT number: Gene cluster number  
 Accession: Genbank accession numbers

| Pkey   | CAT number | Accession   |
|--------|------------|---|
| 300611 | 337193_1   | N75450 AA877636 AW137945 W05248 AA514763 AW972399 AI758397 AW195051   |
| 301187 | 434061_1   | AW976692 AA806542 AA745856  |
| 301254 | 463589_1   | AI049624 AA814705 AW404856 BE078289 BE078292  |
| 301266 | 468223_1   | AA829774 AI082020   |
| 301454 | 534162_2   | AI751738 AA977930   |
| 301615 | 5613_2     | W39477 AK002047 NM_015515 T58707 AA386214 C19007 AA295466 T49621 T47323   |
| 301661 | 7974_1     | AK001735 AF227906 AI815558 AW238991 AL133051 AW272417 AI083492 AI816503 AW888717 AA333166 AI925832<br>BE048352 BE048415 AI141922 AW805674 AW805578 AA633581 AA424632 R71439 AW020988 AW976735 AA883247<br>W37208 AI091039 AW317020 BE221788 AA502917 AW009024 AI141417 BE349081 AI421443 AI080490 AI003921 AI373690<br>AI379240 AA424587 AA740607 AA972391 AA620797 AW271656 AA400517 AI370902 AI680616 AA757270 AA909500<br>N32107 R43738 AI270464 AI870568 AI085139 AA225666 Z41046 AI767739 AI270546 N56779  |
| 301685 | 326972_1   | W67730 Z44630 AA490699 W67596 W76661 R21207   |
| 301804 | 61_1       | AK001468 AA190315 AA374980 AW961179 AA307782 AA315295 AA347194 AW953073 AW368190 AW368192 AA280772<br>AA251247 N85676 AI215522 AI216389 N87835 R12261 R57094 AI660045 AA347193 R16712 AW119006 N55905 N87768<br>AW900167 AI341261 AI818674 D20285 AI475165 AA300756 R40626 AI122827 AA133250 AI952488 AA970372 AA889845<br>AW069517 AI524385 AA190314 AI673359 AA971105 AI351088 AI872789 AI919056 AI611216 AK001472 BE568761<br>AA581004   |
| 301976 | 128835_1   | T97905 AA101672   |
| 302245 | 9482_1     | H18835 R47363 AI460004 N31660 AA454774 AA551759 AI417040 AA694490 AA633315 AI344661 AA708532 AA878567<br>AI802702 AI913465 AW001160 AA932133 AI092908 AA026974 AW628573 AA592910 H18836 AI274428 C00675 AK000048<br>BE313619  |
| 302292 | 27735_1    | AF067797 AB013456 NM_001169 AI791955 AW843925 AI732659 AA577625 AW083143 AW138645   |
| 302476 | 31932_3    | AF182294 NM_016200 AL046942 AI354410 AI697029 AI859557 AW188855 AW105437 AI358735 AW000959 AI491813<br>AW023693   |
| 302623 | 9705_1     | AW836724 BE243668 AB019571 H43803   |
| 302626 | 10441_1    | AK001553 AK001951 AB021870 NM_016282 F01168 AA211870 AA078889 AA312979 AL138385 R70844 AA165658<br>AA007279 AA194688 H65871 AA476639 F01095 AA300170 R39487 AA649126 AA193643 AA418300 BE173477 N84408<br>AW024465 AA406255 BE173412 BE173583 BE173470 AW069288 AA372937 BE504414 AA209472 AI262833 AI628359<br>AI458075 AI476266 AA397706 AI768605 AW243125 AI056436 AA838326 AA810651 AI472025 N35912 AA165622 AI985532<br>AI139528 AA626087 W16998 AI632833 AW130827 AW662551 AA731459 AW780188 AI653447 AI694970 AA810662<br>AI199987 AI587402 AI492972 H65872 AI805624 AW194835 AW994874 R70790 AA836506 N53285 F00181 R83595<br>AI290941 AW936750 AW936703 AW936623 AW936785 AW936691 AW936668 AW936713 AW936788 AW936744<br>AW936613 AW936614 AW936665 AW936702 AW936647 AW936643 AW936712 AW936791 AW936624 AW936672<br>AW936754 AW936696 AW936802 AW936792 AW936589 AW936692 AW936645 AW936746 AW936801 AW936748<br>AW936661 AW936612 AW936697 AW936704 AW936695 AW936626 AW936794 AW936629 AW936577 AW936798 T35617<br>AA375943 R29459 AW936717 AA342108 AW963351 Z24876 AW936708 AW374110 AW936586 W20080 AW936752 W31803<br>AA093709 AA431256 AW803610 AA424959 W76607 AA432267 W72009 R70817 AW778851 AA890563 AA194632 AI089644<br>AI373864 AA890333 AI745574 AI095714 AI567507 AI280712 AW864083 AW468991 N48087 AA860500 AA279471<br>AA993680 AA676504 AI360949 AI052134 AI038657 AI439836 AA629147 AA651840 AA435925 AA854457 AW796472<br>AA838729 AA193407 AA302403 AW958003 AA342107 AA639258 AI435811 AA410342 N25790 AA156454 AI539628<br>AI275854 N58849 AI858171 AW338576 W15321 AA418342 AA780577 W04701 AA630452 AW769154 AI274286 N23736<br>BE465020 AI554346 AI920804 AA969728 AW193440 AI368697 AA115096 AA564981 AA630461 N91475 BE464381<br>AA913741 AA757161 Z24907 C00067 AA649290 AI245223 AA363098 AI520754 AA887983 AI273015 AW878871 AW878981<br>AA480455 AA709267 AW959521 AW959523 N90014 N32441 F00193 AA115095 AA147583 W19813 AI333349 AI197937<br>R39488 AW750110 |
| 302655 | 41899_1    | AJ227892 AA338715 BE074475 BE074469 BE074474 AW006182 AW572953 AI831725 AI762923 AI341466 AW449335<br>BE551686 AI692895 AI040410 AI276881 AI891008  |
| 302758 | 24028_3    | AK001841 H40087 H11121 AW408676 N99603 AA984563 H92041 H11226   |
| 302882 | 458_60     | AW403330 AF062097   |

|    |        |           |  |
|----|--------|-----------|--|
| 5  | 302977 | 47403_3   | AW263124 AI925166 AW105732 AA804479 BE621436 AF086399 W79085 W74440 AW992181 AA389686 AA314311<br>AA173955 AA677564 D59895 D60771 AI887733 C14814 AW162193 D81894 AA732538 AW150919 AA748064 AA769465<br>AA708143 BE327613 AA092726 AI692476 T35673 Z33600 AA134036 AI671394 AI267461 AW362795 AI769759 AA909042<br>AA130042 AW156938 AI753129 AI246205 AI823883 AI752836 D60770 AI336386 AI584003 AW627976 AI348676 D59894<br>AI969795 AW073259 AI400534 AI081318 AI082427 BE550515   |
|    | 302980 | 47495_1   | AI925740 AF086489 W93435 W93345 AA337166 AW966214 AA336257 T11355 AW842435   |
|    | 303011 | 41689_1   | AF090405 AF090407 AF090406   |
|    | 303037 | 35681_1   | AF118395 NM_014317 AW376657 AW848189 AI261617 AI963829 AW848591 AW848598 AW376696 AW848523 AW848450<br>AW848655 AW848183 AW848550 AW376675 AI632752 AI590245 AI431824 AI857990 AI953341 AA888092 AW364968<br>AI188545 AI217741 AW275906 AI311481 AI991404 AI364963 AA628392 AA927982 AW150563 AA503063 AW079470<br>AW512180 AA889371 AW390132 AW609052 AW390112 AW581780   |
|    | 303042 | 5058_1    | AW505345 AF129532 AF126028 AA852108 BE169359 R83701 Z43904 BE613543 AA283163 AA905463 AW067849 R13544<br>R12337 R14020 H98970 AI474918 N56139 AL135669 AW067702 AW372065 AW631389 AA083416 AA287511 AA602923<br>AA488914 AI167215 AW946829 R82855 AI948792 AA371333 AW953883 AW956152 C02539 AA298280 AI932587 AA022742<br>AI983021 AA195252 N58991 R78733 AW083996 H39614 AI365249 AW615389 AI927744 AI089971 N52205 AA083417<br>BE326666 BE349514 AI743785 AI640148 AI378211 AW181881 AI949484 W31374 AW628233 AA418406 AW068010<br>AI708085 AI092696 AI089823 AI277828 AA022660 AI440527 AW054937 AW474104 AI017436 AI159819 AI356716<br>AW473140 AW316518 N34522 AI675092 AI866697 AA864593 AW511185 AA488844 AA904975 N49111 Z39951 R37265<br>AI141362 T25856 R20664 F03163 AI767927 AA805942 D79905 AI914645 AW190553 AI934213 AI458796 AA195385 R82854<br>W31965  |
| 10 | 303072 | 4654_1    | AI566718 AF157833 NM_012133 AI202415 AK002086 AF207598 AI214562 AI202184 AI865579 AA603481 AA483808<br>AA909166 AA774034 AW748102 AW176026 AW351472 BE164787 AA970983 BE622521 BE389817 AW366336 AW366328<br>AW366327 AW366329 AW366335 AW366337 BE269711 T11249 T11264 BE253295 BE256412 BE250882 BE255440<br>BE257663  |
|    | 303149 | 97393_1   | AW963315 AA312995 AA037152 AA088607 AA064770 BE088067  |
|    | 303306 | 11887_1   | AB037732 AW503898 AA215297 BE547488 AW177355 AA046224 AA361664 AA773328 AW512704 AI283330 AI307357<br>AI138263 AA046116 AI219874 AA315431 AW169999 AA492006 AW298002 AA043140 AA131781 AA292383 AA031721<br>AA027867 R31381 AW023352 AI686186 AW467416 AA493914 AA483019 AA483081 AA040871 AA558288 AW070397<br>AW572828 AA693439 AW206584 AA761354 AA907254 AI671019 BE221791 AI915828 AA744724 AA027815 AA131769<br>AA031641 AA837286 AA737401 AI765196 AW086076 AW873024 AI567164 AA744556 AA888910 AI572276  |
|    | 303443 | 224022_1  | AA320525 AW025411 AI684617 AI653685  |
|    | 303502 | 325188_1  | BE174240 AA488528 AL042253   |
| 15 | 303582 | 647662_1  | AA377444 AI458965  |
|    | 303610 | 226089_1  | BE247299 AA323288 AW966142 AA334916 AL046572 BE145095 AW751265   |
|    | 303642 | 284260_1  | AW299459 AA417112  |
|    | 303777 | 244977_1  | AA348491 BE246984 AW505247   |
|    | 303839 | 1770217_1 | Z45939 T54414 T06550   |
| 20 | 303874 | 5013_1    | AL050333 F08138 Z43325 H13393 AA258921 AA224232 BE439918 AL050018 AW363692 AA236615 AA746291 Z19312<br>AA428674 Z28579 T32527 AW952956 R59046 AA403173 AA403171 AW023058 AA461143 BE149531 AA428185 AI382812<br>H42659 AA406086 L48858 AW630177 Z24777 AW675297 AI393859 AI743022 AI669354 AW803015 AA401255 AI952901<br>AW043840 AI808787 AI140662 AA194627 AI140997 AA007454 AA007318 AI469859 AI540581 C06482 AI277356 AI458423<br>AA460839 AA861452 AI080197 AA630781 AA845367 AI125582 AA411705 AA970524 AA699910 AW804640 AW805007<br>AA724226 AI128207 AI696852 AW673064 AA748404 AW771788 AW088185 AI026976 AI537560 AA224233 T24024 T50208<br>AI827319 R17235 T11904 AI816830 R41845 AA639828 Z42124 AA258158 H06057 F02752   |
|    | 303886 | 81595_1   | AW365963 BE141537 BE141538 T19123 R57434 Z43870 AA298099 AA298004 AW963314 AI627790 AA298160<br>BE501485 AW271198 AA195563 AA195584 H28868 AA004370 Z42582 R21338  |
|    | 303929 |           | AW470753   |
|    | 304084 |           | T91986   |
|    | 304143 | 30606_-14 | R88737   |
| 25 | 304165 |           | H73265   |
|    | 304183 |           | H91161   |
|    | 304236 |           | W93278   |
|    | 304350 |           | AA186871   |
|    | 304439 |           | AA398882   |
| 30 | 304495 |           | AA446448   |
|    | 304518 |           | AA461438   |
|    | 304521 |           | AA464716   |
|    | 304546 | 14011_1   | AF113676 X01683 K01396 M11465 NM_000295 K02212 J02619 AB004044 H60588 T72131 T74637 T70970 T73183 T62154<br>R93629 H50855 H80585 H78044 T69186 R95698 H59327 T54018 H83071 R99626 H89864 H91798 T72841 T71108 T72812<br>T54005 T50896 H56102 W01486 W01669 AA137076 H90340 T61854 T61840 N93360 T61844 N53576 T55852 X02920<br>H56350 T58720 H56351 R92748 H56914 H59279 H50665 H56928 T69144 H80448 R91066 H77829 R92479 T55014 T52174<br>T67613 AA845231 H95664 V00496 M26123 AA484470 AI114833 T58716 T64752 T50876 T67858 H48675 T51161 T70409<br>T61715 T72289 H51854 H72171 T50834 H81483 T72132 T58792 T51179 T72833 R29662 T60563 W23562 H94193 T55017<br>T74830 H78469 H90811 T61303 H75867 T71527 T68360 R91065 R91079 T71172 H52923 T50871 T86567 H94691 T69226<br>AV649730 T46850 T56587 T46849 T60552 BE043578 BE042051 T72296 T61001 T58918 T52107 H82324 H47453 R06725<br>D16856 T48282 T52250 R92117 AI287339 H73203 AA318670 D17206 H66626 T69268 H73485 R93078 H73533 R87097<br>T71529 AA885254 AA486074 R94242 H74033 T73643 D12131 AV655901 AA345387 W07278 AW371443 AW371484<br>AW371427 AW371435 AW371467 AW371474 AW371471 AW371453 AW371448 AW371468 AW371486 AW371521<br>AW371463 AW371485 AW371460 T55283 H58030 H68955 AA360298 H73996 T58919 N94213 T50911 H51045 T56077<br>T72988 T71819 H60101 T72439 T68079 T73548 D11609 T61008 H67597 N49781 T73190 H50843 T73140 H61124<br>AV657376 N39304 AW075086 AI247165 H73123 H60258 AA343450 X17122 AW470706 AA336649 AW392737 H75576<br>N76963 T64227 AI803294 W73727 AV649563 AI307406 AW075080 D11525 AI032826 |
|    | 304547 |           | AA486189   |
| 35 | 304636 |           | AA524031   |
|    | 304677 | 2822_15   | AF119850 AI133021 BE561919 BE617370 BE621149 BE561526 BE616541 BE561473 BE618997 BE549320 BE515105<br>AA471325 BE616284 BE559668 BE547034 BE548383 BE299097 BE545965 BE614687 BE615783 BE561441 BE562777   |

BE515345 BE614972 BE560948 BE561499 BE615787 BE267297 BE545452 BE300347 BE514432 BE615122 BE561421  
 BE560145 BE618939 BE544951 BE395272 BE619057 BE396645 BE563309 BE544527 BE408111 BE615925 BE254022  
 BE549220 BE562856 BE279501 BE616173 AA486917 BE560738 BE547802 BE548127 BE561779 BE314412 BE542474  
 BE249907 R35879 BE561536 BE544036 BE549153 BE560885 BE545111 BE256731 BE548221 BE619339 AA130747  
 BE259251 BE294371 BE546522 BE548227 BE612845 BE543601 N87541 H45619 BE547387 BE259822 W25818 BE542131  
 BE547778 BE292868 AW580223 BE513458 T60201 AW364101 BE293099 BE296447 BE270504 AW794068 BE266461  
 BE617287 BE277701 BE265741 BE267576 BE293068 BE265316 BE253499 BE261049 BE513754 BE252038 AA652484  
 BE255820 BE171777 BE393701 BE253247 BE564673 BE261164 AA308384 BE314208 BE293686 BE393327 BE263931  
 BE019750 AW937098 BE315499 AA079590 AA247598 BE387154 AW601680 AA152060 BE253235 BE258730 BE076040  
 BE549252 W05061 NM\_001404 X63526 BE314801 BE397574 AA353735 BE622279 BE561832 BE256022 BE297838  
 BE542783 AA346189 Z11531 AA159202 N27603 H68019 BE616749 BE249954 AA206881 BE619172 AL079889 T50884  
 T73176 H84175 N42762 AA190934 AA199914 BE541310 BE267209 BE304598 BE256955 BE257013 AA213850 BE546575  
 BE410825 BE257518 BE304451 BE392426 BE249978 AW938636 BE280611 BE269434 AA018801 BE249981 AA356501  
 BE546691 BE543217 BE535759 BE548149 BE274370 BE298124 BE621502 BE548075 BE562746 BE383394 BE262657  
 BE544530 BE619215 AI904614 BE546044 BE256949 BE560570 BE538936 BE548558 BE269263 BE296993 BE260919  
 AW862760 AA011045 BE563467 BE396985 AA934616 BE378952 AA100761 BE251832 BE254887 BE545870 BE396771  
 BE543862 BE538869 BE266240 BE542730 BE383556 BE502067 BE552014 AW602866 BE261720 AA016106 BE258671  
 BE281263 AW882888 W56843 AA090219 AW883193 BE618734 BE539981 BE616532 AA338063 BE618675 M55409  
 BE267147 BE545495 BE622854 BE618741 BE299535 BE512709 BE619549 BE621826 BE619650 BE271084 BE304968  
 BE620704 F00658 BE278020 BE543196 AA359668 AW959325 BE296029 BE256156 BE613346 BE617836 BE304900  
 AI142782 BE619732 BE621566 BE396956 AW887816 BE168703 AI219282 BE613404 BE088705 AI379660 AI620510  
 AA659784 W79449 AI767892 AA908586 AW863561 AI142797 BE278846 AA299822 AW954154 AA826854 W32788  
 AW801176 BE257998 BE569129 AA167694 H95000 AI431695 AA613987 T53824 BE514817 AA227510 AA877445  
 AW882770 AA507451 AW820977 BE174939 AI200307 AI445023 AI569411 AA548618 AA724243 AA653140 AA522801  
 AW578153 AA177136 AW192960 R76536 AA548071 AW969645 N86444 AI273891 R49511 AA860967 AA916058 AA826867  
 H25586 BE548884 AI911255 AA199842 AA115835 AA976259 H69325 AA055300 AI474601 W60531 AI719943 H26633  
 H67377 AW410942 AA133585 AA069276 BE260996 AA668262 AA172175 AW973378 AA084253 AA989530 AI250640  
 BE168951 BE564839 AW440725 BE564468 AA070688 BE315049 AA224349 F31228 AW517818 AA523027 AA190811  
 AA161089 AA088923 AI468923 AA129110 AA205266 AI220801 AA181780 AA121359 AA483551 AA501345 AA081217  
 AA865545 AA151850 AA604041 AA583670 AA807148 AA056615 AA486229 AA206889 R43764 AA187051 N88700 T49757  
 AW247901 AA582333 AW194553 AW387155 AA086260 AW166834 AA535771 AA001541 AI275743 AW589611 BE539505  
 AI472623 AA724240 AI270077 R26806 AA018915 AW979353 AI702378 AW591470 AW015803 AW265338 F27619 F37352  
 AA548741  
 AA577816  
 AA582081  
 AA584540  
 NM\_005534 U05875 BE244633 AW163768 AW163746 AI879080 U05877 BE274096 BE390182 BE409310 BE302866  
 AA303965 W31129 AA496185 F07599 Z43842 R24838 N29198 AW877047 AW877121 AW877125 AW363656 AW877123  
 AA377595 BE207637 W20297 AA315151 W35239 T89516 AA058408 AW408191 W23815 AA448929 AI793330 AA173035  
 AI567425 C02282 AW602173 R73773 AI870491 AA287175 AI140367 AI025534 AA370611 AI824138 AI300591 AI300537  
 H04055 AA039282 AI921705 AA738464 AA376968 AW474501 AW157734 AW190741 AW474549 AI590825 BE302836  
 AA128344 AW151252 AI813665 AW514266 W78209 AI368371 AI419996 AA834439 AA085743 AA779622 AW264599  
 BE313816 AA470375 AA572697 AI970559 AI445034 N91836 AI160431 W72709 AI248975 AA757418 AI719633 BE300540  
 AI811310 AI368446 AI619458 AI186860 AA037665 AI126764 AI985833 AA579211 AI126756 AI140467 AA470989 AI188863  
 AA913322 AA825360 AA848066 AA683483 BE206448 AW008445 AA741327 AI022347 AA740593 AW117190 AI906053  
 AA843509 AI139173 AI733305 AI492747 N50093 AA862085 D20407 AI094789 AA449440 AA719906 T89696 AA838546  
 AI140848 AA503840 AI580923 AA828526 AI970569 AA993118 AA082836 AA825533 AA699833 AI263685 AA922853  
 AI160301 AI078729 AA570145 BE046096 AI371141 AI436611 AI186016 AI088757 AI141470 AI187126 AI342299 AI373836  
 AI827816 AI201090 AA970120 AA551892 AA287086 AW841869 AI492741 AI073355 AI192343 AI344429 AI879541 AI423770  
 N94310 AI051206 AI344504 AA854553 AI126700 AA775124 AW572702 W19288 AA748746 AI344385 H38212 AI130719  
 AA626341 AW771310 AW019997 AW771346 AI922093 AA775114 R44542 AA703616 AI362485 AW072333 AI186652  
 AW675704 AI636453 AA709431 R43128 AA588595 AA832391 AI865833 R73681 AA515501 W235701 N94608 AI798771  
 BE045450 AI369476 AI886680 AW197390 Z39908 H03368 AI766761 T28478 AI424369 F03847 N90677 F32611 BE041313  
 AI420363 AW009689 AI933460 AA614856 AI401046 AA056737 AI620906 AI797411 AA039629 BE243267 W23784 AI311585  
 W70312 AA081965 AA643695 AA541686 AI249750 AW298157 AA939276 N63266 AW000886 AA678042 F19175 AI499969  
 AI924421 T25501 Z20802 AI972829 AI826621 AW891571 Z20801 W26292 AW190699 AI640652 AA970137 W77882  
 AI567514 AI570317 D45596 AA905904  
 AA612612  
 AA613107  
 AA618044  
 AA621203  
 AA641638  
 AA642622  
 AA668563  
 AA679005  
 AA679225  
 T81656 AI174932 T82847 T19334 BE563893 BE563101 BE407270 BE386893 AW075527 BE616551 AI879322 AI929728  
 AI557023 AA657885 AI207449 H74068 AA135820 BE560404 BE250148 BE619362 BE612714 BE276295 X55715  
 NM\_001005 T71701 BE513229 BE385149 BE387298 R27196 T55650 AI546980 BE276020 AA483994 AI525170 BE384198  
 AA158297 BE206964 AA259269 H51936 S42658 AA187098 AA186388 AA205135 AA533926 BE083862 BE563679 N88835  
 AA096117 AA225077 BE543870 U14991 U14992 U14990 W40297 W17363 AA377786 AA199796 H14178 H74025 W05492  
 AA187571 BE537410 BE270902 BE277109 BE515193 AW674485 AI433617 AW797409 BE297553 T52590 BE264322  
 BE407417 AW583518 BE392397 AA564507 AA187843 BE620435 BE274727 BE544676 BE559616 BE407160 T59911  
 BE083719 AA375543 N58722 BE562491 BE616109 AA876152 AI439321 AA570592 AA381913 AA381480 AA381930  
 AA381960 AA381565 AA381447 AA381519 AA381183 AA381267 AI453809 AA652710 BE378907 AW276979 AA211342  
 T48442 T56312 N42678 AA134859 BE538136 AA166796 AI698339 AI679955 AI679515 BE551797 T56162 BE409734

BE397723 BE410190 N54738 AA176837 AI200414 AW005899 T90127 BE383057 AI193324 AW265248 AA563942  
 AW970872 AI110677 R66333 AF063551 AI433082 AW162201 AA533266 H46197 AI347056 H80563 R38464 BE378960  
 AA642986 H51899 AA578725 AW172802 BE043720 BE041229 BE043205 AI345540 AW301354 BE041779 BE043807  
 5 AW302305 AW301424 AW268315 AA588176 BE171107 AI312967 AI148696 AI307710 AW272066 AW302329 AI054237  
 AI054050 AA588383 T39250 AI546935 AI174280 AI589974 AI246958 AW301874 AI306015 AI284741 AI349605 AI307731  
 AI340494 AI340550 AI251655 AI370084 AI224265 AW271029 AI224705 AI343155 AI224236 AA076493 BE392932 BE387057  
 AI537940 AA738136 AA744193 AI933432 BE043117 BE392795 AW452224 AA860693 AI193219 AI273441 AW748113  
 10 AW883829 N84473 AW249041 AW970861 AA079174 T50227 AW576854 AW883209 AA420780 AA076320 AW883211  
 AW883853 AW836337 AW883470 AW883331 AA156171 H26617 N90099 AI371291 AA181064 AI286302 T52479 R94956  
 AA678887 AA158113 T54355 N71149 AA181675 H71939 AI308306 BE620398 AA523799 AA094727 AW411258 AW381687  
 AA494005 T48633 AA976344 AI472484 AI500002 AI497712 T70949 AI005589 AA555014 T49791 AA679895 AW572711  
 AI345573 AA541349 AA885550 N69050 AA664708 AI348824 AW072863 AW303049 AI784034 BE138654 BE139281  
 15 AW072723 AI312363 AW572720 AA177047 AW327859 AA713813 AI908759 AI474689 AI890275 AW606076 AA523658  
 AA523366 AA522640 AA340801 AW268340 AA730765 BE349872 AI014737 AW300665 AW392212 AI040590 AA046630  
 AA045165 AI972954 BE550191 AA044995 AI191439 AI191437  
 305295  
 305296  
 305309 15215\_2  
 20 N25117 BE567374 AA188062 R76474 BE394149 AA346106 BE268871 W32999 AW793468 AI075912 D53484 AW512628  
 AA829843 AA831923 AA187162 AI375868 AA576000 U41448 NM\_001029 X69654 AW068471 AI262598 AI719560 N24138  
 AA588388 F22286 AA385512 AA155919 AI347783 F22561 AI016457 AI708902 AA622797 AW084953 W35385 AA845254  
 AA683406 AA773365 AA443835 AI626039 AA715725 AW900657 AW900718 AW148625 AA187280 AA583906 AA916044  
 25 AI500256 N75427 N87626 AA921883 N89069 AW327607 AI377311 AI735263 AI039117 H51501 AA384041 AA908565  
 AI022990 AA530944 AW169092 AA640128 F31984 AI421320 AA902400 T47242 AA187283 AA676731 BE466262 AI358868  
 AI829098 AI720169 AA602834 AA075051 AI016333 AI832618 AI626112 AW028704 AI807675 AI554601 AI735043  
 AW275957 AA135844 AA102517 AI202054 AA836808 BE536365 H71341 AI719486 AI497920 AI581707 AA601142 AI626130  
 AA991422 AI344639 AA740155 AI350162 AW780123 AI056774 AI057132 AA532368 AI524528 AA720807 AI553829  
 AI688278 R76475 AA136006 AA614598 AA224772 AA181050 AI278337 AA583702 AA932358 AI708568 AA657866  
 30 BE173137 BE174600 AA181944 T48413 AA716293 H87639 T86323 T86324 AI885811 N35961 AA186975 AI221285  
 AI707796 AW273582 AA775295 AI719122 AA091156 H23786 AA225182 AA779355 AI718506 AA148325 T47243 AA358519  
 AA001821 F28492 AA737274 F25722 AI038467 H51521 AA576106 F35015 AA548697 BE073926 AA748127 T48414 D54761  
 AA863282 AW384185 AA086265 AA508410 AW297113 AA342512 AI888475 T55298 AA758944 AI719674 AA501428  
 AI718704 BE165424 AA961799 AA962669 AA983161 BE173686 AA886117 AA872983 AW793184 AA665189 T55130  
 AA613898 H91152 AW797007 AA071280 AA071377 AI880677 AA659794 F35193 C15194 T51001 T70211 AI625921  
 35 AA908590 AA551606 T69882 AA077153 AA346105 AA225274 AA358552 F18375 AW002113 D55051 AW368652 AW368778  
 AI721092 AI335641 AI202753 T81175 AI720731 AA747418 AA304776 AW022506 AI299308 T25496 AI749286 AI460263  
 D53639 AI124989 AA593610 T68867 AI460138 AA358484 AW975054 AA321580 F28232 AA353257 AA303644 AW068878  
 AW068730 F25874 AA300373 D56049 F26553 BE568636 AA385266 AA316385 AA303470 X77770 AA090283 AW327251  
 40 AA316866 F37411 BE385312 BE253018 AA320071 AA308568 D51956 T27782 BE393178 D51473 AA364881 D52094  
 AA303753 AA314074 AA369588 H91202 D52058 BE393600 BE389958 F33433 BE382989 F28039 H51520 AW239409  
 F32141 BE304844 AA316043 F30160 AA353239 F19450 F19208 AA187556 AA961772 F31331 AA075050 AA187320  
 AI374585 F27581 F32548 AA308515 T51089 AW964109 AA128451 BE386435 H51500 AA570763 AA224794 AA157613  
 T68942 AA096244 AA186484 AA360942 W78176 AA147724 AA699717 AI052202 AA095483 D52120 T25967 AW880072  
 AA903320 AA082357 H71393  
 45 AA706676  
 AA714040  
 AA714180  
 AA782331  
 AA804508  
 50 AA815428  
 AA826884  
 AA828156  
 AA829363  
 AA835250  
 55 AA889992  
 AA902309  
 AA906983  
 AA908609  
 AA916314  
 60 AA932299  
 AA936312  
 AA937901  
 AA946870  
 AA954198  
 AA961821  
 65 24593\_27 BE271766 AF116719 BE513150 H49670 BE397881 BE614531 BE562280 BE387230 BE270556 AA333494 U43901  
 AW630924 U43901 BE620882 X15005 AI630308 AA224482 BE265468 BE566970 BE542957 BE567127 BE564341  
 BE257276 N85360 R47201 AA130008 BE393122 T57407 BE304440 BE538285 BE543610 W04320 BE261521 BE395193  
 BE277139 BE513224 BE542480 BE298611 BE542479 BE545346 BE546436 T70965 BE620480 J03799 BE397891 BE562251  
 70 AA373403 T55389 BE393268 BE562409 AA324299 AA317559 BE393578 AA064690 BE562188 AI525344 AW672959  
 BE561902 R19651 BE616792 AA308270 BE563600 N31380 T74744 T58212 N31805 BE542361 AA357300 BE543041  
 BE560993 N76928 BE389448 BE262375 BE538256 R96424 R20398 BE615985 S37431 T72413 NM\_002295 BE281623  
 AA076198 AI205269 AA334045 T55370 AW800811 T57971 AA134763 AA209447 R19067 AA053200 BE542788 BE299054  
 BE268607 N49627 T51676 BE561411 BE619391 AL031311 BE388868 BE279523 BE560315 BE271784 BE253962  
 75 BE619010 BE258112 BE382896 BE535693 BE620389 AA204831 BE267276 BE547775 BE249951 BE537902 BE298296  
 BE537758 BE612439 BE267227 BE295417 H81474 AA083251 BE545562 BE279950 AW389264 BE545054 BE294352  
 BE382915 BE259414 BE542092 BE298128 BE273820 BE271588 BE253890 BE563184 BE546606 BE263148 BE614968

BE542402 H67612 BE560882 BE546130 BE543147 BE540497 BE537524 BE613753 BE547997 N50015 BE614047  
 BE618974 H94126 AA113031 BE546168 AA102514 BE565376 BE543414 BE272689 BE545988 BE537457 BE617006  
 T86726 BE512888 X61156 BE560027 BE538834 AA076331 T59466 BE559529 BE257047 T59764 AA191701 AA385844  
 BE562585 AA366436 AA335298 BE019631 BE378577 BE535521 BE396621 BE279958 BE549122 BE537798 AI338812  
 AA961833 H21616 W04906 AA928220 BE548328 AA089878 T71120 T56899 BE542669 AA082136 T82381 T94025 H81332  
 AA772988 BE545437 R56691 H93222 H70111 AA090016 H55990 AI184974 BE620968 BE539322 AA160696 AA346768  
 BE389278 N86754 AA092482 AA086091 N83880 AA147808 AA147849 BE544762 AA342408 AA204832 AA147998 R28783  
 AA328498 H79896 R91237 AA301536 AA206450 AA358599 N84489 AA094379 R29091 AA206157 AA489490 AW889947  
 AA167032 W19474 AI241644 AI432647 AA665658 AI718549 AA737795 AI207622 AA599406 AW514994 W35312 T59884  
 AA643137 AI200347 AW591918 T57822 AI190984 T59412 AA658064 AW973080 AI370110 BE176259 AA730904 AA872869  
 AW869253 AA227371 AA968861 AI281565 M14199 AI266346 T55798 H99950 W23712 AI033135 T96423 AA128434 R97053  
 AI266269 AA114134 AI264933 AA593164 R99334 R44109 R71185 H81360 AA532483 R56847 H66334 AA083252 AA135211  
 R44437 AA320995 AA983284 T96327 AA622534 D51200 AI469759 N89876 AA129995 AI168036 R43516 AA835456  
 AA508702 AA989565 AA552538 AA491480 T56856 R58891 AA112376 H75933 AA099846 AI553658 AI538107 AA197339  
 W40255 AI358440 AA506867 AW176499 N89447 AA548318 AW089158 AW800311 AA205200 H47125 H61801 AI817550  
 AA885896 H61800 AW516691 AW089792 AA206057 AW411349 AW301332  
 306413 AA973288  
 306462 AA983397  
 306469 AA983792  
 20 306501 AA987294  
 306563 AA995296  
 306610 AI000635  
 306723 AI026151  
 306744 AI031882  
 25 306792 AI042426  
 306810 AI057294  
 306820 AI074408  
 306896 AI093383  
 306953 AI124971  
 30 307010 AI140014  
 307012 AI140212  
 307034 AI142526  
 307041 AI144243  
 307056 AI148675  
 35 307064 AI149335  
 307074 23309\_6 BE171526 AI159895 AI362169 AI557175 AI816211 W84586 AA443432 AA478090 BE273860 AW067965 BE616889  
 AI905033 AW161521 W07464 AA402174 T72985 AA086142 AI816076 AI929055 AW404703 X05607 NM\_000099 X52255  
 AI362017 N31169 AW964703 AI346361 AA430624 AI419306 AI089436 AA478163 AI422740 AA405246 AA581187 M78181  
 AA114823 AI569697 AA456462 AI056312 AI394439 AI056388 AI394465 AI419320 AI096478 AA554829 AW003877 AI394482  
 AI149531 AI014323 AI096480 AI816174 AA454916 AI050879 AA723285 AW055076 M27891 X61681 AI564229 AA464371  
 AW026289 AI356750 AI357354 AI368416 AI685243 AI609768 AI816118 AI184254 AI423042 AI357585 AW055080 AA738018  
 AI890978 AA284648 AA994535 AW026440 AI421217 AA379815 AW026352 AI218999 AA713818 AI424074 AI003039  
 AA947130 AW051798 AA371755 AI633210 AI342529 AW007903 AI949112 AA776633 BE379201 AI361017 AI589045  
 AW026349 AI421109 AI361213 AI356896 AI356628 AW027415 AW005034 AW304939 AW175887 AA622325 AI206007  
 AI422765 AI953860 AI055990 AA363573 AW027888 AI498831 AI937335 AA385758 AW880771 AA523539 AA315198  
 AA371942 AI420662 AI418260 AI422729 AI424095 AA335397 W93398 AA335301 AA336083 AI088933 AA933630 AI207592  
 AI422814 AW880787 H46870 AI422104 AI784528 AI422028 AI423492 AI418908 AI476176 AI418491 AI262265 AW054997  
 AI401222 AW935107 R75680 AI356763 AA086053 AA320998 AW959124 AW951566 AA582437 AA618600 AI928791  
 AI141394 AI380094 AI077661 AA570651 AI418220 AA586896 AW051500 AW628321 AW168616 AI423428 AA775816  
 50 AA464260 AA363434 AA808067 AI356585 AA365044 H39961 AI421421 AI084346 AI421590 AI678845 AI401337 T68178  
 AA533499 AI420682 AI422976 AA704329 H24665 AA425987 AI041100 AI423091 AA704378 AA482076 AA688207 AA588778  
 AI582532 AA594350 AI799230 AI817567 AA365706 AA526950 AA325184 AI624278 AA989696 AI148572 AA854613  
 AI499904 AA320404 AW005422 AA533158 AI146970 AA291655 AI149169 BE327985 AW328286 AA922081 AA729897  
 55 AW799758 AA364028 AA843331 F20435 AW001524 AI285253 AA916618 AI422358 AI094747 AI094754 AI262629 AI597587  
 AI281139 AI096859 AI537257 AA846506 AA973472 AA845740 AA977830 AW157016 AI174258 AI460190 AI089868  
 AI095037 AI089864 AA723298 AI074244 AI096812 AA687428 AI422334 AI359247 AA782354 AI460199 AA454568 AI096464  
 AA426217 AI097391 AA434378 AI097359 H93031 AI160095 AI373821 AA477632 AI564398 AI150904 AI346637 AW026849  
 AA573822 AW192676 AA369379 AI094919 AI290053 AA319991 AA627370 AA702191 AA363695 AI096607 AA992885  
 60 AI183336 AI361555 AA890199 AW768597 AI421336 AW001458 AI095758 AW027909 AI150989 AA620294 AA974335  
 AI306477 AI146532 AW103696 AW025355 AA588754 AA772569 T51534 AI077588 AI051136 H45665 AA641555 AA335249  
 AI359662 AA335173 AI160230 AI359451 AI096516 AI832913 AA856626 AW129193 AI057277 AI299349 AI801789 AI360811  
 AW026246 AW009786 AI146563 H54066 AI141357 AI290397 AI359760 AI073390 AA369438 AA782546 AI360379 AW516525  
 AW026180 R73226 AI057139 AA918124 AA652886 AI359602 AW008584 AI361599 AA858102 AI361575 AI360860  
 65 AI363088 AI095068 W74631 AI096419 AI359287 AA617644 AI027185 AI721209 AA579636 AI535703 AI535742 AA411125  
 AI535680 BE206595 AI076856 AI084271 AI356202 AI076863 AW104687 F22288 F37131 AI135467 AA722050 AI283930  
 AW193140 AA584226 T68253 AI362193 W46392 AI419269 AA533764 AA565340 N29369 AI282034 AI074482 AA464433  
 AA335520 AI816542 AI311529 AA402731 AI361536 AA936821 AI090604 AI278884 AI272311 AA454842 AW511753  
 AW316946 AA577547 AI625298 AW008545 BE206455 F35252 AA478251 AA894702 AW055036 AA318997 AI051266  
 70 AA916075 AA468716 AW575608 AI097241 AA478250 AI813393 AA961776 AW168580 AI362156 H56144 AI351645  
 AW051887 AA464434 AA782238 AW069713 AW328287 AI361335 H22417 H28467 AA774593 H69881 AA948632 AA987632  
 AW026296 AW069809 AI270651 AA290771 AW027914 AA402315 W02152 H23679 F34337 AI367943 AI422932 AA464435  
 AA630257 AA630372 AA216227 AI354596 AA725375 AI832369 AI864731 AA642707 AI143086 N90262 AI471066 AI475343  
 AA478008 N81104 AA838768 AI147250 AA195967 AI097264 AI148176 AI041946 H64634 AA304727 AI245016 AA947243  
 75 AA320023 AA411983 AI285077 AI005504 AA602808 AA954547 AA844184 AW900840 AI356806 AA934005 AA658077  
 R54958 T29906 AI983056 AA364947 F27923 N69886 AI418676 AI149599 AI418632 AI636479 AA114824 AW168203  
 AA854263 AA290780 AW008659 AW149229 H18245 AA559148 AW008735 AI088875 W84605 AI149397 AW515564

|    |        |  |
|----|--------|--|
|    |        | AA722497 AI422059 AI394499 AI475192 AI589945 H24677 AA548688 AI130975 AI130862 F30448 AA313754 AI148377<br>AA996252 AI915211 AA886378 AI371520 AI879469 F24100 F28104 AA503929 F19709 H28541 AA993350 AA320635<br>AI783926 D11806 D11707 T60062 H62085 F36915 H19496 AA558311 AA485384 AA037016 AA435523 T60162 AA602795<br>AA477963 AA335919 AA335626 AW008995 AA991946 AI554752 AI370939 AI400912 AA250788 AA320310 AA999770<br>AI313034   |
| 5  | 307084 | AI160527   |
|    | 307095 | AI167910   |
|    | 307156 | AI186762   |
|    | 307243 | AI199957   |
| 10 | 307264 | AI202211   |
|    | 307288 | AI205169   |
|    | 307391 | AI225058   |
|    | 307542 | AI280859   |
|    | 307565 | AI282468   |
| 15 | 307586 | 9205_7<br>AW367324 AA513829 BE300206 AA134200 AA126689 AW975636 AA375190 AA229666 BE562156 AW957307 AI133468<br>BE562075 BE293171 H43213 AA095435 AW405074 N28598 AA579221 AA096022 AA315868 AI557910 AA096172<br>BE616496 R83253 NM_006013 M64241 BE386761 W03711 BE261517 BE298374 BE278491 BE255200 BE304448<br>AA206687 BE386934 BE618525 T49742 AA532479 H92515 BE299890 N84711 W80584 W04247 AV654899 BE313651<br>AA361932 BE616507 BE568551 H18202 W93773 S35960 M81806 AA352773 BE281085 H69766 N87825 AA090275 M73791<br>AA090218 H79923 H85682 N89270 AI001784 N83949 BE567780 R73099 H63621 AA714817 BE536488 BE407585<br>BE394061 BE619086 T71244 BE563254 C04286 BE293916 T67271 BE544286 BE538828 S64169 BE254604 BE295847<br>BE407999 BE616137 BE618994 AL121934 AA084264 AA084260 BE615364 BE541737 R30878 BE537874 BE293082<br>BE622544 BE280736 BE560886 BE512983 BE512970 AI469402 BE279821 BE615519 BE563424 AI027614 S35959 T67270<br>AA377416 H20002 W04857 AB019572 H91139 BE279255 W02004 W05580 AA524537 AW970052 H91379 AI223104<br>AI219245 AI190973 R73100 AA618062 AA583590 F37254 AA054426 AW951771 AI582983 AI535778 AW183619 AI285499<br>AI205618 R18051 T50234 BE252578 AI025348 AA738325 AI190989 AW732196 AA531213 AI439364 AA292351 N70489<br>AA126581 AI567584 AI264944 C14996 AI351267 W65440 H68843 AA548110 H60635 T50114 AA187603 AA404627<br>AA614381 AA835418 AA192999 T49333 AI246998 AA070742 AW388661 T49876 AI865171 BE261134 T50166 AW752163<br>AI866246 AI497721 AA022533 AW388584 C05870 AW388551 AW375685 AA079408 AW388685 AW388565 AW388289<br>AA346577 AW752158 AW752161 AW752157 AW752160 AW388563 AL031276 AW375687 AA528266 AW578215 BE250977<br>H66751 AA054338 AA084261 AW972379 AA084265 AA446308 AL046840 AW580988 T50626 AA634027 AA492347 D51630<br>BE261889 AW581036 AA302508 AB007170 AW752162 AW375688 AW375709 AW752152 AI499767 AA936623 AW455210<br>AW388395 AW388383 AW581024 AW388313 AW797015 T25776 AA503342 BE378698 AA501932 D17268 AA493357<br>AA932910 AW075203 BE171178 AW388379 AW388583 AW388335 AW581005 AA369640 AW388419 AW388324 AW579474<br>AW388588 AW388719 AA582262 AW388360   |
|    | 307612 | AI290787   |
|    | 307626 | AI300035   |
|    | 307640 | AI301992   |
|    | 307646 | AI302236   |
| 40 | 307783 | 697809_1<br>AI347274 AW844024  |
|    | 307856 | AI366158   |
|    | 307954 | AI419692   |
|    | 308047 | AI459633   |
|    | 308050 | 9482_1<br>H18835 R47363 AI460004 N31660 AA454774 AA551759 AI417040 AA694490 AA633315 AI344661 AA708532 AA878567<br>AI802702 AI913465 AW001160 AA932133 AI092908 AA026974 AW628573 AA592910 H18836 AI274428 C00675 AK000048<br>BE313619   |
|    | 308106 | AI476803   |
|    | 308139 | AI494477   |
|    | 308186 | 33276_16<br>T81656 AI174932 T82847 T19334 BE563893 BE563101 BE407270 BE386893 AW075527 BE616551 AI879322 AI929728<br>AI557023 AA657885 AI207449 H74068 AA135820 BE560404 BE250148 BE619362 BE612714 BE276295 X55715<br>NM_001005 T71701 BE513229 BE385149 BE387298 R27196 T55650 AI546980 BE276020 AA483994 AI525170 BE384198<br>AA158297 BE206964 AA259269 H51936 S42658 AA187098 AA186388 AA205135 AA533926 BE083862 BE563679 N88835<br>AA096117 AA225077 BE543870 U14991 U14992 U14990 W40297 W17363 AA377786 AA199796 H14178 H74025 W05492<br>AA187571 BE537410 BE270902 BE277109 BE515193 AW674485 AI433617 AW797409 BE297553 T52590 BE264322<br>BE407417 AW583518 BE392397 AA564507 AA187843 BE620435 BE274727 BE544676 BE559616 BE407160 T59911<br>BE083719 AA375543 N58722 BE562491 BE616109 AA876152 AI439321 AA570592 AA381913 AA381480 AA381930<br>AA381960 AA381565 AA381447 AA381519 AA381183 AA381267 AI453809 AA652710 BE378907 AW276979 AA211342<br>T48442 T56312 N42678 AA134859 BE538136 AA166796 AI698339 AI679955 AI679515 BE551797 T56162 BE409734<br>BE397723 BE410190 N54738 AA176837 AI200414 AW005899 T90127 BE383057 AI193324 AW265248 AA563942<br>AW970872 AI110677 R66333 AF063551 AI433082 AW162201 AA533266 H46197 AI347056 H80563 R38464 BE378960<br>AA642986 H51899 AA578725 AW172802 BE043720 BE041229 BE043205 AI345540 AW301354 BE041779 BE043807<br>AW302305 AW301424 AW268315 AA588176 BE171107 AI312967 AI148696 AI307710 AW272066 AW302329 AI054237<br>AI054050 AA588383 T39250 AI546935 AI174280 AI589974 AI246958 AW301874 AI306015 AI284741 AI349605 AI307731<br>AI340494 AI340550 AI251655 AI370084 AI224265 AW271029 AI224705 AI343155 AI224236 AA076493 BE392932 BE387057<br>AI537940 AA738136 AA744193 AI933432 BE043117 BE392795 AW452224 AA860693 AI193219 AI273441 AW748113<br>AW883829 N84473 AW249041 AW970861 AA079174 T50227 AW576854 AW883209 AA420780 AA076320 AW883211<br>AW883853 AW836337 AW883470 AW883331 AA156171 H26617 N90099 AI371291 AA181064 AI286302 T52479 R94956<br>AA678887 AA158113 T54355 N71149 AA181675 H71939 AI308306 BE620398 AA523799 AA094727 AW411258 AW381687<br>AA494005 T48633 AA976344 AI472484 AI500002 AI497712 T70949 AI005589 AA555014 T49791 AA679895 AW572711<br>AI345573 AA541349 AA885550 N69050 AA664708 AI348824 AW072863 AW303049 AI784034 BE138654 BE139281<br>AW072723 AI312363 AW572720 AA177047 AW327859 AA713813 AI908759 AI474689 AI890275 AW606076 AA523658<br>AA523366 AA522640 AA340801 AW268340 AA730765 BE349872 AI014737 AW300665 AW392212 AI040590 AA046630<br>AA045165 AI972954 BE550191 AA044995 AI191439 AI191437 |
|    | 308191 | AI538878   |
| 75 | 308243 | AI560037   |
|    | 308249 | AI560998   |

|    |        |           |  |
|----|--------|-----------|--|
|    | 308256 |           | AI565498   |
|    | 308337 | 3328_20   | J00124 AA158848 AW452899 NM_000526 AW384139 BE140708 BE140690 H28274 BE140404 AW844922 BE157342<br>AW580043 AW384049 BE157333 BE157179 H44127 BE140671 BE614947 AW380727 BE184532 AA808075 R73201<br>AW949948 AW950242 BE183306 BE183297 AA587174 BE183243 AI591112 AI609229 AI590690 AW606369 AI609238<br>AI608848 AI590885 AW084944 AI609220 AI609210 AA583890 AI609274 AI609226 AI608878 AA550914 AW170430<br>AI608947 AI608606 R88146 M28646 AA587248 AA583985 AW872988 AA583993 AA586980 AI284481 AA584051 AA587024<br>AW800015 AA587313 AI283394 AI143314 AA595434 AA583917 AA922332 AA584021 AA583576 AA862004 AW873168<br>AW796060 AA862049 AW873161 AW079705 AI366769 AA583724 AA583889 D29065 AA113288 AW083705 AA584334<br>AA586711 AA583558 AW117878 R72295 AA583842 AW996598 R73138 H28224 R88223 AA586985 AW369022 AA602252<br>AA641063 R72845 AA861952 AA587149 AA587308 AW369033 AA584344 AW799955 AA583732 AW368990 AA862083<br>D29595 R53012 AA587254 D29462 AA158849 R72649 BE181541 D29604 AI127631 D29297 H44051 AA583575 AA586762<br>D29122 AA583878 AA595946 AW368880 AA587318 AW368881 |
| 5  |        |           | AI611044   |
|    | 308354 |           | AI611044   |
|    | 308362 | 792518_1  | AW998989 AI613519  |
| 10 |        |           | AI623988   |
|    | 308380 |           | AI669859   |
|    | 308457 |           | AI671311 BE501055 AW813781 AW813651 AW836924   |
|    | 308462 | 816335_1  | AI679292   |
|    | 308484 |           | AI708573   |
| 15 |        |           | AK000142 AW243187 AI738593 AW505395 BE009209   |
|    | 308578 |           | AI758754   |
|    | 308615 | 33893_1   | AI761036   |
|    | 308667 |           | AI811707   |
|    | 308676 |           | AI863908   |
| 20 |        |           | AI869642   |
|    | 308781 |           | AI879831   |
|    | 308928 |           | AI880489   |
| 25 |        |           | AI972052   |
|    | 308957 |           | AW081820   |
|    | 308991 |           | AW169657   |
|    | 309000 |           | AW193825   |
|    | 309243 |           | AW293999   |
|    | 309398 |           | AW294725   |
| 30 |        |           | AW296076   |
|    | 309582 |           | AW419122   |
|    | 309634 |           | AW451663   |
|    | 309828 |           | AW302773 AW303087 AI254651   |
|    | 309830 |           | T51842 T51888 T51099 T51008  |
| 35 |        |           | R23313 R23323 Z25059 AA359123 AW965886 BE167187 AI808503   |
|    | 309839 |           | AA330047 AW962512 AI983896 AI590126 AA837373 AI638125 AW139338 AA765161 AW194751 AA765897 AA261818   |
|    | 309939 |           | F13656   |
|    | 309977 |           | BE385785 T69279 AA703325 AA678398  |
| 40 |        |           | T87431 AW300867 T87330 AA515973 AI242970   |
|    | 310341 | 651442_1  | AA524394 AW015969 AA158731 AI831401 AI955800 AW272596 AW272595 AI281799 R34231 R35558 AW378527   |
|    | 310985 | 114598_1  | AI523875 R45782 R45781   |
|    | 311395 | 252948_1  | AW368634 AI702169 AI245179 AW368646 BE545574 AA249018 AW368633 N27553  |
|    | 311736 | 193321_1  | N49730 N52659  |
|    |        |           | N59555 AA362113 AA362136   |
| 45 |        |           | BE396243 AA040818 AA134134 AA040155  |
|    | 311965 | 404771_1  | AA284333 AW468119 AA284334 AA810992  |
|    | 312129 | 338029_1  | AA079229 AA079201 AA078874   |
|    | 312339 | 151127_1  | AA215470 AA215547 AW470551   |
|    | 312405 | 765247_1  | AW752247 AA167125 AA161495   |
|    | 312996 | 187327_1  | AA827082 AA732246 AA167611 AA830741  |
|    | 313078 | 1653847_1 | AA740616 AA654854 AA229923   |
|    | 313139 | 255206_1  | AW971564 AA251226 AA835824 AA748600  |
| 50 |        |           | AA535749 AA574021 AA302428   |
|    | 313561 | 98898_1   | AI791249 AA452038 AW817263 AW817051  |
|    | 313603 | 199797_1  | AA533447 AF158241 AI240598   |
|    | 313781 | 116301_1  | AA737345 AA682286 AI799378   |
|    | 313825 | 173444_1  | AW105663 AA693880 AW517398 AI768507 BE220851 AW978538 AA831489   |
|    | 314033 | 153186_1  | AA789249 AA904217 AA904142   |
| 55 |        |           | AA838114 AW629478 AA883713 AI620552  |
|    | 314043 | 155125_1  | N58024 T58194 T11693 N64222 T05848   |
|    | 314138 | 179960_1  | AW868221 T11286 H53526 AW630814 AA381202   |
|    | 314183 | 31294_3   | AW992356 D82445 D82303 T11347 AW995841 AW995909 AA133068 AW859596 AW859595 AA644624 AW631466   |
|    | 314367 | 210554_1  | AI903549 AI903561 AI380423 AW364047 AW373615 AW373755 AA634799 AW860107 AA336073 AW960638  |
|    | 314703 | 303778_1  | AI284099 AI284098 AI201463 AI972977 AA224831 BE183312 AI567336 AW389340 H39906 AW770710 AA411736 AI763378  |
| 60 |        |           | AW340972 AW516306 AI400359 AI745530 AW511407 AA406542 AA524982 AI628029 AW008882 AI858439 AW074041   |
|    | 315021 | 344728_1  | AI379597 AA132792 AW189012 AI538874 AA857364 AA224830 AW590976 AI678604 AW512052 AA593133 AI142902   |
|    | 315858 | 406384_1  | AI992380 AA888921 AW517569 AI679729 AA577041 AI640743 AI282492 AA494400 AW074288 AA551421 AA904079   |
|    | 316055 | 409389_1  | AA505483 AW513312 AI469669 AI872908 AI290109 AI610272 AA829570 AI611723  |
|    | 316652 | 454745_1  | AW407564 AA262049 W19405 AA504733 T12641 AW973724 AA558125 AW993087 T12640 T11447 AA521285 AI820042  |
| 65 |        |           | AA457028 AI674737 AI688648 N92749 AI439620 AI218005 AA731321 AA828303 T12590 AI685371 AA504636 AA825573  |
|    | 316897 | 474090_1  | AW172614 AA291292 AA649919 AW576505 AA262152 H29888 AA830757 AA910433 AA808481 AI971807 AI767874   |
|    | 317617 | 1660859_1 | AI216422 AA831483 AI566351 H41305 AA879438 AA807123  |
| 70 |        |           |  |
|    | 318171 | 269002_1  |  |
|    | 318175 | 139564_1  |  |
| 75 |        |           |  |
|    | 318230 | 193526_1  |  |

|    |        |           |   |          |  |
|----|--------|-----------|---|----------|--|
|    |        |           | 318581  | 162718_1 | AI633044 AW016212 AW241143 AA769058 R43272 AW068958 AA210918 AA293774 AI748815 |
|    |        |           | AI763294 AI333114 AI277384 AI088297 AI468477 AI824624 AW189606 AI631751 Z40749 AI984673 AI671316 AA189024   |          |  |
|    |        |           | AW235412 BE178426 R24677 R40635 H05100 R40597   |          |  |
| 5  | 318601 | 366837_1  | AA975853 AI915867 AW341040 AI985422 AI733001 AW300448 AA865842 AI913757 AI283772 AI928188 AA593410          |          |  |
|    |        |           | AW771149 AI732663 T39921  |          |  |
|    | 318647 | 767064_1  | AI557774 AI526159 AI526153 AI541503 AI541531 AI526149 AI526018 AI526177 AI557786 AI526140 AI557743 AI541492 |          |  |
|    | 318652 | 73477_1   | T53260 AI659829 AI620887 BE619759   |          |  |
|    | 318721 | 250974_1  | Z28504 H85748 AA418060 AA356371 AW965781 W45671   |          |  |
| 10 | 318740 | 15117_1   | NM_002543 AF079167 AB010710 AF035776 AA114093 AJ131757 R07930 AI243883 R62556 AA682386 AA620341 R78357      |          |  |
|    |        |           | R31345 H67054 AI300045 BE174919 AW082681 N75085 AI928113 R23598 AA044216                                    |          |  |
|    | 318975 | 458663_1  | Z44110 F13225 R13941 BE542680 T75468 N36522 BE207328 AW136340 AA808805 AI633843 AW875667 AI637851           |          |  |
|    |        |           | AI990628 AW608757 AI636889 AW136314 R58192  |          |  |
|    | 319128 | 273945_1  | H09041 H10254 AA968912 Z45974 AA393820  |          |  |
| 15 | 319141 | 1533572_1 | F12377 Z46143 F05137 T74119   |          |  |
|    | 319198 | 1535512_1 | F07354 R11946 H16080  |          |  |
|    | 319436 | 1596398_1 | T70298 H58072 R02750  |          |  |
|    | 319478 | 765461_1  | AI524124 R06841 R06842  |          |  |
| 20 | 319488 | 10887_1   | AK001330 AA356435 BE313393 BE293644 BE251929 AA808340 BE409475 AA331948 N91096 AW402232 AW402994            |          |  |
|    |        |           | AW386322 AA004739 AA459479 AW579400 W68758 AW673556 BE313041 AA455698 AL045680 AA134589 AW606254            |          |  |
|    |        |           | BE301261 AW976697 AW968467 AW976703 W68453 AA004688 AW976701 AA223724 AA565953 AA215565 AA744555            |          |  |
|    |        |           | AW193840 AI086227 AW970769 BE300513 AI458782 AI183406 AI309531 AA455644 AA128908 AA588705 AI138389          |          |  |
|    |        |           | AI476292 AA515291 AA524425 AA459254 AA600279 AA614836 AA769786 AA492544 AL045681 AA765178 AI864425          |          |  |
|    |        |           | AW780369 BE246640 AA926793 AW054669 N63744 AA206610 AA729135 AA766112 AI553635                              |          |  |
| 25 | 319551 | 357371_1  | AA761668 AA573621 R92814 R09670   |          |  |
|    | 319599 | 798430_1  | H56112 H58047 AI630710 N58742   |          |  |
|    | 319808 | 7069_3    | T58960 AA609180 AA621130 AI927236 AA431075  |          |  |
|    | 319827 | 776890_1  | T62926 AI565392 T62778  |          |  |
|    | 319884 | 171016_1  | AA325579 AW961004 BE004785 T73234 AA209403 N54886   |          |  |
| 30 | 319936 | 57653_1   | W22152 AV647377 AV647331 AA320693 T79025 F23202   |          |  |
|    | 319983 | 1814568_1 | T81429 T95572 T95563  |          |  |
|    | 320001 | 331474_1  | AA873350 T82429 T82428  |          |  |
|    | 320445 | 132254_1  | R26830 R33029 AA115761 AW118148 AI743741 AA954284 AI934165 AI088310 AI123759 AW340232 AI089180 AA700861     |          |  |
|    |        |           | AI129973 AI088552 AW963119 AA359516 R33916  |          |  |
| 35 | 320488 | 368456_1  | AI817336 R32883 AA595590 AI743065 R31386  |          |  |
|    | 320503 | 25164_1   | NM_005897 AF156857 AA346876 BE545147 AI003306 N45644 AW889728 BE007236                                      |          |  |
|    | 320661 | 159_1     | AB030034 BE304778 NM_016653 AF251441 AA307684 AW957882 AF238255 NM_016595 R57782 AA244505 AA864846          |          |  |
|    |        |           | AW601475 AA232750 AI417539 AA232253 AW294490 AA626441 AW814670 AW814669 N95341 AI123874 AA100075            |          |  |
|    |        |           | AW337275 AW804295 AI922069 BE161875 AI470677 AI242841 AA402558 AI435815 AA402496 AI359093 AA505991          |          |  |
| 40 | 320727 | 36759_63  | AW197200 AA234622 AA258509 H17033 AI799498 AI263346 AA236466 AA258354 N24807 R14272 AA100160                |          |  |
|    |        |           | AW003360 AI971548 AA017585 X80306 X91133 AJ276100 X91132 AF064499 AF064495 AF063768 X83713 AW951310         |          |  |
|    |        |           | AW975565 AA721610 AA715972  |          |  |
|    | 320746 | 29157_2   | AW975814 AA282765 AA811755 AA731129 BE219297 AA128302 R71285 AA095218                                       |          |  |
|    | 320789 | 252516_1  | R78712 AA603646 R78713  |          |  |
| 45 | 320809 | 254953_1  | W16480 AA376361 N83837 R81853 AA361779  |          |  |
|    | 320825 | 29807_1   | NM_004751 AF102542 AW360893 AF038650 BE304708 AW360892 AW360931 AW842622 AA307800 BE292814 AW582119         |          |  |
|    |        |           | AW582122 AW374998 AW374874 AI587061 AA583339 AW662377 AW192901 AW887756 AW887761 AI955582 AI150400          |          |  |
|    |        |           | AA568218 AA583146 AI832775 AA294858 AI445680  |          |  |
|    |        |           | D59968 D81035 C15620 D80887 D81432 C15618 D60320 D80661   |          |  |
| 50 | 320858 | 1509963_1 | D62269 AW022615   |          |  |
|    | 320909 | 996651_1  | AA737314 AA682280 AA010792 AA143573 AA953433 AA745273 AA188649 AA011221                                     |          |  |
|    | 321012 | 84379_2   | AA610649 AI699484 H59558  |          |  |
|    | 321253 | 375160_1  | BE245833 BE539992 AI380940 AW952644 AA535470 R84610   |          |  |
|    | 321286 | 236538_2  | AI696519 BE464779 AW296343 BE550149 AW470402 AA129660 T78937 AA342648 N71662 H82431 AI302712 AV660681       |          |  |
|    | 321307 | 69583_1   | R85409 AA962323 AI680732 AA889147 AA932629 AW103527   |          |  |
| 55 | 321354 | 116028_2  | AA078493  |          |  |
|    | 321370 | 41905_1   | AJ227900 AI094933 AW051119 F00947   |          |  |
|    | 321412 | 624592_1  | AI674383 AI865710 AI201451 AI659387 U25919 BE093109 AW366305 BE141926 BE141913 AW854334 AW854342            |          |  |
|    |        |           | BE141916  |          |  |
| 60 | 321424 | 107996_1  | D59886 AA779752 AI655936 AW976526 AA235034 AA744353 T26888 AA235103 R96569 AA057301 AA057286 D61635         |          |  |
|    |        |           | H87227  |          |  |
|    | 321467 | 43034_1   | X13075 X13076   |          |  |
|    | 321524 | 45508_1   | R39382 BE467537 AI657156 AI375103 AW021134 AI479241 BE326541 AW150836 AI684065 R35463 AA678409 AI694321     |          |  |
|    |        |           | AW470057 AW608873 N62359 AI702778 AI701838 AI655208 BE465196 R51845 R38307 AW393336 AW043913 AA782285       |          |  |
|    |        |           | AI205974 L13824 L23311 AI635429 L13826  |          |  |
| 65 | 321632 | 286374_1  | AW812795 AA419617 H87827 AW299775 AW382168 AW382133 BE171659 AW392392 BE171641 AA541393                     |          |  |
|    | 321828 | 44964_2   | R59890 R60548 N64863 AI224545 N69114 AI811204 AW518902 AI184866 AI440169 AA809472 H63089 AW952971           |          |  |
|    |        |           | AW337382 AI872923 N73882 AA334161 AI537113 H63731 AI383952 N41701   |          |  |
|    |        |           | N78520 AW606984 AI287235 AA973956 N49122  |          |  |
| 70 | 321875 | 430904_2  | AW089866 N63915   |          |  |
|    | 321920 | 1023421_1 | N76794 N94221 W04156 AW897535   |          |  |
|    | 321974 | 1280819_1 | AL137517 BE072492 AI127076 AW196207 AW294979  |          |  |
|    | 322035 | 33334_1   | AF085833 R69689 AW341677 AA923375 BE327566 AW630415 R69601 AW615339   |          |  |
|    | 322092 | 46678_1   | AF075083 H52291 H52528  |          |  |
|    | 322136 | 46802_1   | AF085975 H53458 H53459  |          |  |
| 75 | 322175 | 46877_1   | AI357412 AI870708 AI590539 W07459   |          |  |
|    | 322303 | 704603_1  | W76622 AF086372 W72660  |          |  |
|    | 322309 | 47372_1   |   |          |  |

|    |        |          |   |
|----|--------|----------|---|
|    | 322331 | 47467_1  | AF086467 W81444 W81445  |
|    | 322340 | 47509_1  | AF088076 W95222 W92523  |
|    | 322394 | 27492_1  | AW068287 AA310079 BE336702 AA356318 AA306059 AA346785 AW402633 AA311210 AW402909 N76879 AW402913            |
| 5  |        |          | AW401920 AA321636 AA354474 C17297 C16938 AA311774 M29871 NM_002872 Z82188 AW405674 H94176 R89281            |
|    |        |          | AA214723 AI014482 AW949347 T27749 AW804226 AW796964 AW404581 AF077208 NM_014029 W68830 W79652               |
|    |        |          | AA353375 AW575218 AA552192 AA521232 AA702695 AA033975 AW407827 AA829948 N94402 AW628604 AI523308            |
|    |        |          | N57605 AA641662 H42477 N52784 AI753478 AA768493 AA845729 W47391 N55270 AI090117 R89282 BE206172             |
|    |        |          | AA076650 AA595650 AI218931 BE049397 AI433110 W74114 H94277 AI358627 AI085221 AI862818 AA835967 AW103905     |
| 10 |        |          | AI640644 AA835507 AA856887 AA694392 AW337542 AI524410 BE045500 AI440060 AI358801 AW028238 AW205248          |
|    |        |          | AI718264 R48618 AA357358 AI695002 AA897549 AW081065 AI433360 AI810783 AI620963 Z82188 AA360224              |
|    | 322518 | 38914_1  | AI133446 T50819 AF147343 T50665   |
|    | 322567 | 39354_1  | AF155108 AW877241 AW393512 BE160738 AW384889 AW610272 BE160915 BE160774 BE160744 AW836696 AW384919          |
|    |        |          | AW836739 BE160743 BE160814 AW610275 BE160965 AW580785 AA662739 BE160941 AW821136 AW592083 AA449860          |
|    |        |          | AI798661 AA310698 AW302768 AW268932 AW268741 AI250559 AW302879 AW821181 AW580243 AW384922 AW606812          |
| 15 |        |          | AI345641 AW821143 AW384873  |
|    | 322610 | 21773_1  | BE242847 AA159840 NM_016216 AF180919 BE262663 BE312610 W53026 BE093965 BE004620 AW992549 AA069408           |
|    |        |          | R66803 BE002445 T80130 N67797 AA765401 AA765829 AW837997 AW837993 AW838011 AW838012 AW837996                |
|    |        |          | AA069435 W52118 AI457469 AA954977 R39354  |
|    | 322694 | 35627_1  | NM_014125 AF090919 AF075371 AI110872 BE070571   |
| 20 | 322735 | 91819_1  | AA086123 AA026296 AA086041 AA026297   |
|    | 322817 | 449438_1 | AA777274 AI761381 AI738617 C02420   |
|    | 322890 | 117967_2 | AA649792 AA640427   |
|    | 322933 | 127352_1 | AA099759 AA100511 AA687172  |
|    | 322944 | 130324_1 | AA112573 AA112574 AA984323  |
| 25 | 322959 | 132817_1 | AI267606 AA121045 AA126521  |
|    | 322968 | 17218_2  | AI272141 AI879676 AF070669 W25179 AA534016 AA533386 AA010740 AF124147 W16493 W56636 AA258911 AA321677       |
|    |        |          | H44503 AA642777 AA081800 W69885 H82507 AA536128 AA326782 AA326783 AA353693 AA354642 R73311 AA354400         |
|    |        |          | W79820 W16502 AA301647 AI202303 AA453926 AA705795 AA011128 AA929033 AI393389 AA845133 AI445640 AI677727     |
| 30 |        |          | AI818296 AI369820 AI539292 AI870541 W69797 AI871096 BE550803 N35853 AA644019 N27809 R49769 AA738197         |
|    |        |          | AI565700 AW207656 AA587216 AA669237 AI906947 AI809956 AI740905 AW043811 AW182476 AA659844 AI742797          |
|    |        |          | AI832103 AI660967 R86125 AI674667 AI808074 AI869284 AI336214 AI218002 AI338629 AI857930 AW183986 BE044333   |
|    |        |          | AW135467 AI826077 AI357643 AI475486 AA478855 AW172550 AI553942 AA868731 AW268850 AI123793 AI887022          |
|    |        |          | AA046935 AI361954 AI091737 AI682235 AI367076 AI088882 AI088682 AI312679 AA046955 AW027546 AI660019 AI696174 |
| 35 |        |          | AW008626 AI268337 AI568959 AW027409 AI040014 AW134559 AA479953 AA910082 AI301458 AW028352 AI017863          |
|    |        |          | AI268915 AI185866 AI265907 AI274195 AW051540 AW027515 AI380435 AA883117 AI279396 AA846628 AW628235          |
|    |        |          | AW206201 AW628510 AA954276 AI301405 AI827185 AI553978 AI200301 AI470343 AA933953 AI914937 AI362849          |
|    |        |          | AW085066 AI204021 AA631192 AI351701 AA748663 AA993806 AA580146 AW027744 AA580016 AA897344 AI042638          |
|    |        |          | AI473196 AA995065 AW027720 AI217421 AI935604 AW449411 AW237094 AI653348 NM_003107 X70683 AI470473           |
| 40 |        |          | AI765137 AI193479 AI253050 AI470510 AI399828 AI371461 AI185518 N20940 R49816 N79977 W56599 N24649 W78113    |
|    |        |          | N78761 AI817673 AI911482 AW205984 AI240186 AI828016 AI942449 X65661 AW751587 AI392808 AI624192 AI950969     |
|    |        |          | AA573260 AI203361 AI479942 BE041834 AW305351 AI918327 BE048713 AW071712 BE041565 AI139260 BE466360          |
|    |        |          | BE502737 AW007819 AW071887 AI742130 AI344020 AW772112 AI932275 AI992189 AI197801 BE219990 AI990863          |
|    |        |          | AI536934 AI336275 AI971955 AI798204 AI870429 AI652390 AI080187 BE219486 AI185434 AW628564 AW072399 AI656370 |
| 45 |        |          | AI498606 BE041559 AI743591 AW515805 AI087833 AI917506 AI123191 AI858043 AI334046 AI242585 AI636670 AI919478 |
|    |        |          | AW771487 AI417185 AI468527 AW137861 AI554782 AI130733 AW005164 AI910551 AI189135 AI963934 AI985482 AI660396 |
|    |        |          | AI497963 AW204662 AW137602 AI382505 AI493485 AI185987 AI078841 AI830054 AI378223 AI351299 AA937301          |
|    |        |          | AA242817 AA258359 AW027603 AI935204 AI500360 AI569741 BE551058 AW275536 AI457854 AI142093 AW028288          |
|    |        |          | AI286002 AI279114 AI364121 AI341323 AI190436 AW002607 AI242488 AI338122 AI368600 AI340276 AI417994 AI190234 |
| 50 |        |          | AI275527 AI934886 AI498274 AI813630 AI075339 AI087976 AI459251 AI989477 AW004046 AI992190 AI885279 AI479475 |
|    |        |          | AI698030 AI473294 AI951648 AI699587 AI660602 AI873018 AW613987 AI808297 AW270159 AW572955 AW195908          |
|    |        |          | AW469034 AW197100 AA885164 AW611668 AI143038 AI910560 AA418374 AW3341092 AI871169 AI937136 AI204003         |
|    |        |          | AA775707 AW590759 AW593350 AW572981 AI197905 AI660941 AI743469 AW237017 AI808587 AI984962 AA418254          |
|    |        |          | AI828104 AA625231 AI832151 H84232 AI240215 AI911775 AI219668 AI336801 AA232630 AI343471 W69129 N93602       |
| 55 |        |          | AA768883 W04386 AI086277 AA983433 W07646 AA458584 N86625 AI384055 AI928089 W25479 AA242952 AI763303         |
|    |        |          | AI225039 AI740896 AA953758 W69240 AA558331 AI760593 AA558712 AW992121 AW992157 W69115 BE328596 AI953190     |
|    |        |          | W95311 AI950195 AI739605 AI857262 W69185 AA884586 AI198104 AI127451 AA905932 AA723310 AI936623 AA732940     |
|    |        |          | AI332918 AI221396 AI336095 AI200067 AI824853 D55893 D52697 D56205 AA232764 T53299 H84555 AA076539 AA158347  |
|    |        |          | BE298430 AL134493 AW732398 AW750740 AW578208 N36572 AA453861 AA252914 AA234197 AW576988 AW577034            |
| 60 |        |          | AA025199 AW577052 AW385538 AW576996 AW577021 T83230 AA421529 AI918492 AA909038 AA507030 AA654561            |
|    |        |          | AA064597 AW001594 AW469192 AI368002 AI142435 AW379382 W93438 AA076387 AI802344 AI097013 AA987215            |
|    |        |          | AA635282 W93349 AI017818 AA421564 AA158348 AI140004 AA506259 AW473184 AA236350 AI138669 H96873 AA974889     |
|    |        |          | AA643735 AA995463 AA995471 AA809555 AA253225 AI298682 AI572515 T53300 AA064596 AA193589 AA025118            |
|    |        |          | AI669682 AA610638 T90774 AI972332 AI280776 T27980 AW136058 BE000428 AI378691 AA961520 BE049142 AI1311424    |
| 65 |        |          | AA283211 AI344071 AI344007 AI344097 AI582410 AL036314 AW798038 AI905228 C15325 AA380386 AW958417            |
|    |        |          | AW630531 BE538239 T70488 AA088296 T34175 T31626 D54331 D53142 AA029415 AW946823 AI914128 AA355446           |
|    |        |          | T34322 BE006559 M85677 AA034335 T31463 AW804007 AA256591 D55128 AI535884 D55192 N23605 T31802 AA326899      |
|    |        |          | AW999156 AA355201 AW999306 AI091590 BE172021 AA029490 BE000255 AW339939 AW150093 AI872098 AI274876          |
|    |        |          | T06303 AA857909 N23606 AA922714 AI914104 AI285281 AW999919 AI339803 AI081354 AA972184 AI049566 AW151583     |
|    |        |          | AI682455 AA088257 AI217050 BE551774 AI277033 AI252627 AA910406 AI369422 H46634 AI873113 AA033710 AW078579   |
| 70 |        |          | AI636452 N23010 AA357263 AA256592 T05786 AA884195 AA406145 AA907807 AA482840 AI637691 AA654523 AA911495     |
|    |        |          | T06601 AW594370 AW016524 C15324 AA622519 AA340191 AA174168 C13992 T69433 T96576 AW166822 T96575             |
|    | 323011 | 139750_1 | AA580288 AA315655 AA133031 AA377748   |
|    | 323166 | 162676_1 | AA291001 AA188974 AA290616  |
|    | 323216 | 6526_1   | AA332145 AA331790 AW962563 AA868189 L13837 T34468 AA055882 AA096148 AA092327 H57062 R59098 R11247           |
| 75 |        |          | F07659 Z44949 AF131829 L13835 T79889 AA252451 N28984 H85260 AL046384 AW995631 R58386 AI061651 AW376050      |
|    |        |          | AW379789 W90347 AA450157 AI799939 AA461340 W02347 AA233095 N39675 AA659441 AW995284 W17060 R32252           |

|    |                            |                                   |  |
|----|----------------------------|-----------------------------------|--|
| 5  | 323243                     | 140566_2                          | AI042599 AL046385 AI970370 AA744764 AI249761 AI628106 R32668 AI863011 AI923998 AI186798 N26601 AI141864<br>N34992 AI377031 N23934 AI683466 BE219548 AA622032 AW089867 AA243717 N79547 R59099 AW241293 AI917545<br>AW103697 AI383179 AW517527 AW193642 W90348 AW381409 R11195 AA461166 AA836624 AA280285 AW242055<br>L13836 N89647 L13834 AI358605 AA452023 AI868391 H57063 AW075868 N20590 Z40695 R37603 R28484 AA251913<br>F03914 AA055772 N43752                       |
|    |                            |                                   | W47525 AA134047 BE391212 AA330333 AA376355 BE304871 BE167342 H87402 AA631722 W45724 AA715517 AI925438<br>AI804849 AW241617 AW403807 AI653435 AA134048 AW747874 AI922327 AI814967 AI935895 AA228865 AW504076<br>AA225008 AW673858 C03914  |
|    |                            |                                   | AW675572 AI248270 T85161 AL133848 T70731 T69747  |
| 10 | 323244<br>323333<br>323430 | 647858_1<br>62251_1<br>63341_1    | AV651680 AA228883 AA367341 AW962458 AA628024 AW172426 AI767785 AA313012 AW963323<br>AW062479 AW062488 AW062491 AW062480 AW938564 AW062478 AA322408 AA324351 AW938595 AW938598 BE162389<br>AW176556 AW938599 AW838792 AW938566 BE162305 BE162377 AW938570 AW062459 AW176555 AW938562<br>AW938568 AA251701 BE162320 AW938597   |
|    |                            |                                   | AA287406 AA261844 AA261845 AA287355 AA810895   |
|    |                            |                                   | AA278246 AW292815 AA278703   |
| 15 | 323479<br>323538<br>323632 | 193343_1<br>194627_1<br>217887_1  | AW247696 BE265140 AW403615 AL037647 AA312336   |
|    |                            |                                   | AL041844 AL040002 AL039950   |
|    |                            |                                   | AA323414 AW664013 AI809377 AI276041 AW296883 AI798340  |
| 20 | 323731<br>323753<br>323835 | 226193_1<br>12462_4<br>506747_1   | AK002161 AA327102 AI056868 AI743901 AI139018 AI199114 AI076003   |
|    |                            |                                   | AL042005 AL042006 AA911481   |
|    |                            |                                   | AA347566 AA346521 AI111169   |
| 25 | 323898<br>324048<br>324231 | 243407_1<br>267284_1<br>975669_1  | AA378739 AW964174 AA570564 AI076833 AW265063 AW006805 AA480656 AW004789  |
|    |                            |                                   | W60827 AL079968 AL047234   |
|    |                            |                                   | AA464018 AA464079 AA468142   |
| 30 | 324430<br>324432           | 312113_1<br>312487_1              | AA464510 AA631257 AI740516 AI739132 AW972467 AI741376 AW068935 AI467852 AI752240 AI123717 AI754551<br>AW205510 AW044211 AW028889 AW198033 AI538632 AA513096<br>AW500954 AW501111 AW501394  |
|    |                            |                                   | AW502122 AW502125 AW501663 AW501720  |
|    |                            |                                   | AW502257 AI014241 AA100360 BE298534  |
| 35 | 324512<br>324575<br>324609 | 1155396_1<br>1156071_1<br>65704_1 | AW299534 AW299896 AA504765 AA505099 AA505100 AA584753 AW136415 AA768306<br>BE397649 H14413 BE397689 BE514098 H53372 AA448021 R57944 AI307272 BE259369 H72331 BE251092 T27364<br>AA001666 AA044433 AA875998 AW075405 AW338356 AA001667 AW300173 AW514944 AW468914 AA604673 AA702749<br>AA805550 AA447621 AA934104 AI373527 AA604794 AI911203 AI500644 AI291383 AA731133 BE350633 AA044604 H95689<br>H14366 AV660983 AA912893 AI369587 AI382271 AA917508 AW138391 BE622560 |
|    |                            |                                   | AI376331 AI819150 AI097038 AI351100 AW504689   |
|    |                            |                                   | AW503713 AA352950 AA044972 BE618246 AA335047 AW962269  |
| 40 | 324692<br>324715           | 72231_1<br>351987_1<br>290035_2   | AA557952 AA677593 AA618150   |
|    |                            |                                   | AI739168 AA426249 AI199636 AW505198 AW977291 AA824583 AA883419 AA724079 AI015524 AI377728 AW293682<br>AI928140 AA731438 AI092404 AI085630 AA731340<br>T85872 T48305  |
|    |                            |                                   | AA640770 AI683112 AA913009   |
| 45 | 324728<br>324783<br>324848 | 210991_2<br>389615_1<br>371388_1  | AA602539 D59262 AI684171 N46711 AW021857 D19768  |
|    |                            |                                   | AA613792 AW182329 T05304 AW858385  |
|    |                            |                                   | AK001379 AK001411 AW795711 T06997 AA287540 AA354538 AW957773 AI632268 AI651003 AI689650 AI809332<br>AW304483 AI805269 AA278506 AA862381 AA287875 AW628545 AI085761 AW025965 AI658615 AW628879 AW139496<br>AI214278 AA902745 AA991679 BE540102 AW593658 AI745602 AA744687 AI285441 AA807089 AI218314 AA721449<br>AI202987 AA432129 AI285502 AI281462 AA731319 BE082573  |
| 50 | 325071<br>325176           | 1562044_1<br>700767_1             | H09693 H09699 T09229   |
|    |                            |                                   | T19142 AI351168 T52843 BE241963  |
|    |                            |                                   |  |

**TABLE 1B**

Table 1B shows the genomic positioning for those primekeys in Table 1 that lack unigene ID's and accession numbers. For each predicted exon, the genomic sequence source used for prediction is listed. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique number corresponding to an Eos probeset  
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.  
 Strand: Indicates DNA strand from which exons were predicted.  
 Nt\_position: Indicates nucleotide positions of predicted exons.

|    | Pkey   | Ref               | Strand | Nt_position       |
|----|--------|-------------------|--------|-------------------|
| 15 | 332792 | Dunham, I. et.al. | Plus   | 73381-73768       |
|    | 332908 | Dunham, I. et.al. | Plus   | 1934283-1934366   |
|    | 332909 | Dunham, I. et.al. | Plus   | 1946582-1946735   |
| 20 | 332913 | Dunham, I. et.al. | Plus   | 1963539-1963843   |
|    | 332952 | Dunham, I. et.al. | Plus   | 2472864-2473012   |
|    | 332958 | Dunham, I. et.al. | Plus   | 2516164-2516310   |
|    | 332961 | Dunham, I. et.al. | Plus   | 2521424-2521555   |
|    | 332975 | Dunham, I. et.al. | Plus   | 2599641-2599702   |
| 25 | 332991 | Dunham, I. et.al. | Plus   | 2686938-2687372   |
|    | 333119 | Dunham, I. et.al. | Plus   | 3288316-3288640   |
|    | 333131 | Dunham, I. et.al. | Plus   | 3350064-3350170   |
|    | 333139 | Dunham, I. et.al. | Plus   | 3369495-3369571   |
|    | 333156 | Dunham, I. et.al. | Plus   | 3617584-3617790   |
| 30 | 333222 | Dunham, I. et.al. | Plus   | 3979706-3979803   |
|    | 333254 | Dunham, I. et.al. | Plus   | 2521424-2521555   |
|    | 333348 | Dunham, I. et.al. | Plus   | 4711908-4712181   |
|    | 333349 | Dunham, I. et.al. | Plus   | 4713940-4714084   |
|    | 333366 | Dunham, I. et.al. | Plus   | 4798273-4798469   |
| 35 | 333384 | Dunham, I. et.al. | Plus   | 4907535-4907610   |
|    | 333385 | Dunham, I. et.al. | Plus   | 4907928-4908032   |
|    | 333391 | Dunham, I. et.al. | Plus   | 4916697-4916780   |
|    | 333488 | Dunham, I. et.al. | Plus   | 5396233-5396310   |
|    | 333520 | Dunham, I. et.al. | Plus   | 5586133-5586296   |
| 40 | 333524 | Dunham, I. et.al. | Plus   | 5612620-5612780   |
|    | 333532 | Dunham, I. et.al. | Plus   | 5622804-5622937   |
|    | 333580 | Dunham, I. et.al. | Plus   | 6142935-6143145   |
|    | 333585 | Dunham, I. et.al. | Plus   | 6234778-6234894   |
|    | 333597 | Dunham, I. et.al. | Plus   | 6331421-6331536   |
| 45 | 333619 | Dunham, I. et.al. | Plus   | 6562799-6562926   |
|    | 333671 | Dunham, I. et.al. | Plus   | 7038849-7039193   |
|    | 333680 | Dunham, I. et.al. | Plus   | 7071730-7071794   |
|    | 333682 | Dunham, I. et.al. | Plus   | 7076641-7076760   |
|    | 333763 | Dunham, I. et.al. | Plus   | 7692491-7692630   |
| 50 | 333764 | Dunham, I. et.al. | Plus   | 7693573-7693716   |
|    | 333769 | Dunham, I. et.al. | Plus   | 7696625-7696707   |
|    | 333770 | Dunham, I. et.al. | Plus   | 7700384-7700476   |
|    | 333849 | Dunham, I. et.al. | Plus   | 8018323-8018472   |
|    | 333875 | Dunham, I. et.al. | Plus   | 8135505-8136179   |
| 55 | 333882 | Dunham, I. et.al. | Plus   | 8153002-8153169   |
|    | 333922 | Dunham, I. et.al. | Plus   | 8381385-8381444   |
|    | 333928 | Dunham, I. et.al. | Plus   | 8468844-8469015   |
|    | 333947 | Dunham, I. et.al. | Plus   | 8579888-8579966   |
|    | 333949 | Dunham, I. et.al. | Plus   | 8589634-8589791   |
| 60 | 333968 | Dunham, I. et.al. | Plus   | 8681004-8681241   |
|    | 333983 | Dunham, I. et.al. | Plus   | 8813593-8813668   |
|    | 333995 | Dunham, I. et.al. | Plus   | 8855296-8855424   |
|    | 333997 | Dunham, I. et.al. | Plus   | 8866668-8867255   |
|    | 334003 | Dunham, I. et.al. | Plus   | 8892882-8892970   |
| 65 | 334012 | Dunham, I. et.al. | Plus   | 9007456-9010221   |
|    | 334047 | Dunham, I. et.al. | Plus   | 9428152-9428211   |
|    | 334063 | Dunham, I. et.al. | Plus   | 9731991-9732085   |
|    | 334066 | Dunham, I. et.al. | Plus   | 9739568-9739680   |
|    | 334078 | Dunham, I. et.al. | Plus   | 9809783-9809863   |
| 70 | 334118 | Dunham, I. et.al. | Plus   | 10344273-10344384 |

|    |        |                   |      |                   |
|----|--------|-------------------|------|-------------------|
|    | 334122 | Dunham, I. et.al. | Plus | 10411792-10411901 |
|    | 334150 | Dunham, I. et.al. | Plus | 10529221-10529854 |
|    | 334220 | Dunham, I. et.al. | Plus | 12718720-12718857 |
| 5  | 334298 | Dunham, I. et.al. | Plus | 13424763-13425914 |
|    | 334324 | Dunham, I. et.al. | Plus | 13539210-13539323 |
|    | 334335 | Dunham, I. et.al. | Plus | 13608488-13608705 |
|    | 334433 | Dunham, I. et.al. | Plus | 14273261-14273429 |
|    | 334532 | Dunham, I. et.al. | Plus | 14792798-14792901 |
| 10 | 334561 | Dunham, I. et.al. | Plus | 14987299-14987447 |
|    | 334616 | Dunham, I. et.al. | Plus | 15176123-15176470 |
|    | 334628 | Dunham, I. et.al. | Plus | 15310346-15310415 |
|    | 334630 | Dunham, I. et.al. | Plus | 15322614-15322744 |
|    | 334631 | Dunham, I. et.al. | Plus | 15325949-15326116 |
|    | 334661 | Dunham, I. et.al. | Plus | 15477716-15477786 |
| 15 | 334677 | Dunham, I. et.al. | Plus | 15517449-15517560 |
|    | 334696 | Dunham, I. et.al. | Plus | 15665919-15666002 |
|    | 334714 | Dunham, I. et.al. | Plus | 15760702-15760767 |
|    | 334718 | Dunham, I. et.al. | Plus | 15775491-15775599 |
| 20 | 334720 | Dunham, I. et.al. | Plus | 15792931-15793085 |
|    | 334727 | Dunham, I. et.al. | Plus | 15942616-15942750 |
|    | 334739 | Dunham, I. et.al. | Plus | 16004120-16004225 |
|    | 334740 | Dunham, I. et.al. | Plus | 16009324-16009547 |
|    | 334769 | Dunham, I. et.al. | Plus | 16170704-16170876 |
| 25 | 334872 | Dunham, I. et.al. | Plus | 19162417-19162565 |
|    | 334876 | Dunham, I. et.al. | Plus | 19185336-19185400 |
|    | 334883 | Dunham, I. et.al. | Plus | 19223107-19223253 |
|    | 334891 | Dunham, I. et.al. | Plus | 19299770-19299944 |
|    | 334900 | Dunham, I. et.al. | Plus | 19315678-19315743 |
| 30 | 334902 | Dunham, I. et.al. | Plus | 19317083-19317195 |
|    | 334914 | Dunham, I. et.al. | Plus | 19495158-19495275 |
|    | 334916 | Dunham, I. et.al. | Plus | 19572924-19573846 |
|    | 335044 | Dunham, I. et.al. | Plus | 20842088-20842682 |
|    | 335081 | Dunham, I. et.al. | Plus | 21113871-21113937 |
| 35 | 335158 | Dunham, I. et.al. | Plus | 21569610-21569666 |
|    | 335164 | Dunham, I. et.al. | Plus | 21585912-21586014 |
|    | 335166 | Dunham, I. et.al. | Plus | 21587100-21587213 |
|    | 335170 | Dunham, I. et.al. | Plus | 21623383-21623967 |
|    | 335188 | Dunham, I. et.al. | Plus | 21669118-21669328 |
| 40 | 335189 | Dunham, I. et.al. | Plus | 21673403-21673472 |
|    | 335200 | Dunham, I. et.al. | Plus | 21743499-21743881 |
|    | 335211 | Dunham, I. et.al. | Plus | 21774611-21774680 |
|    | 335219 | Dunham, I. et.al. | Plus | 21875591-21875688 |
|    | 335221 | Dunham, I. et.al. | Plus | 21882840-21882968 |
| 45 | 335225 | Dunham, I. et.al. | Plus | 21890315-21890448 |
|    | 335255 | Dunham, I. et.al. | Plus | 22032258-22032661 |
|    | 335287 | Dunham, I. et.al. | Plus | 22299047-22299299 |
|    | 335361 | Dunham, I. et.al. | Plus | 22807292-22807445 |
|    | 335364 | Dunham, I. et.al. | Plus | 22833430-22833586 |
| 50 | 335369 | Dunham, I. et.al. | Plus | 22843392-22843506 |
|    | 335468 | Dunham, I. et.al. | Plus | 23787245-23787367 |
|    | 335481 | Dunham, I. et.al. | Plus | 24082522-24084870 |
|    | 335488 | Dunham, I. et.al. | Plus | 24118744-24118839 |
|    | 335496 | Dunham, I. et.al. | Plus | 24164386-24164545 |
| 55 | 335497 | Dunham, I. et.al. | Plus | 24167666-24167869 |
|    | 335499 | Dunham, I. et.al. | Plus | 24176698-24176869 |
|    | 335504 | Dunham, I. et.al. | Plus | 24182110-24182199 |
|    | 335599 | Dunham, I. et.al. | Plus | 25043628-25043775 |
|    | 335623 | Dunham, I. et.al. | Plus | 25138489-25138547 |
| 60 | 335653 | Dunham, I. et.al. | Plus | 25329710-25329802 |
|    | 335687 | Dunham, I. et.al. | Plus | 25445952-25446064 |
|    | 335690 | Dunham, I. et.al. | Plus | 25455442-25455625 |
|    | 335692 | Dunham, I. et.al. | Plus | 25468557-25468725 |
|    | 335697 | Dunham, I. et.al. | Plus | 25481456-25481649 |
| 65 | 335701 | Dunham, I. et.al. | Plus | 25513366-25513807 |
|    | 335708 | Dunham, I. et.al. | Plus | 25541777-25541907 |
|    | 335739 | Dunham, I. et.al. | Plus | 25698550-25698826 |
|    | 335742 | Dunham, I. et.al. | Plus | 25712654-25712771 |
|    | 336003 | Dunham, I. et.al. | Plus | 28406289-28406759 |
| 70 | 336015 | Dunham, I. et.al. | Plus | 28640586-28640673 |
|    | 336016 | Dunham, I. et.al. | Plus | 28646816-28646947 |
|    | 336018 | Dunham, I. et.al. | Plus | 28660880-28660978 |
|    | 336019 | Dunham, I. et.al. | Plus | 28663992-28664102 |
|    | 336020 | Dunham, I. et.al. | Plus | 28683778-28683851 |
|    | 336021 | Dunham, I. et.al. | Plus | 28686482-28686559 |
| 75 | 336023 | Dunham, I. et.al. | Plus | 28698240-28698343 |

|    |        |                   |      |                   |
|----|--------|-------------------|------|-------------------|
|    | 336071 | Dunham, I. et.al. | Plus | 29264457-29264684 |
|    | 336090 | Dunham, I. et.al. | Plus | 29413020-29413162 |
|    | 336107 | Dunham, I. et.al. | Plus | 29987731-29987869 |
| 5  | 336121 | Dunham, I. et.al. | Plus | 30048054-30048129 |
|    | 336124 | Dunham, I. et.al. | Plus | 30053441-30053500 |
|    | 336132 | Dunham, I. et.al. | Plus | 30107247-30107412 |
|    | 336135 | Dunham, I. et.al. | Plus | 30123235-30123335 |
|    | 336194 | Dunham, I. et.al. | Plus | 30443138-30443282 |
| 10 | 336235 | Dunham, I. et.al. | Plus | 31122315-31122623 |
|    | 336367 | Dunham, I. et.al. | Plus | 33942937-33943058 |
|    | 336379 | Dunham, I. et.al. | Plus | 33995071-33995243 |
|    | 336439 | Dunham, I. et.al. | Plus | 34186130-34186215 |
|    | 336502 | Dunham, I. et.al. | Plus | 34268953-34269083 |
| 15 | 336572 | Dunham, I. et.al. | Plus | 34446383-34446496 |
|    | 336602 | Dunham, I. et.al. | Plus | 13424060-13424582 |
|    | 336721 | Dunham, I. et.al. | Plus | 3371522-3371586   |
|    | 336739 | Dunham, I. et.al. | Plus | 2599641-2599702   |
|    | 336766 | Dunham, I. et.al. | Plus | 4905608-4905684   |
| 20 | 336833 | Dunham, I. et.al. | Plus | 6856506-6856634   |
|    | 336836 | Dunham, I. et.al. | Plus | 7077262-7077326   |
|    | 336878 | Dunham, I. et.al. | Plus | 9200300-9200399   |
|    | 336880 | Dunham, I. et.al. | Plus | 9250034-9250123   |
|    | 336902 | Dunham, I. et.al. | Plus | 10385555-10386053 |
| 25 | 336917 | Dunham, I. et.al. | Plus | 11228329-11228403 |
|    | 336919 | Dunham, I. et.al. | Plus | 11351181-11351274 |
|    | 336924 | Dunham, I. et.al. | Plus | 11525273-11525527 |
|    | 336946 | Dunham, I. et.al. | Plus | 12337073-12337258 |
|    | 336953 | Dunham, I. et.al. | Plus | 12988791-12988889 |
| 30 | 336979 | Dunham, I. et.al. | Plus | 14270748-14270816 |
|    | 337169 | Dunham, I. et.al. | Plus | 23529987-23530214 |
|    | 337175 | Dunham, I. et.al. | Plus | 23782209-23782374 |
|    | 337182 | Dunham, I. et.al. | Plus | 23934889-23934962 |
|    | 337238 | Dunham, I. et.al. | Plus | 27141465-27141776 |
| 35 | 337303 | Dunham, I. et.al. | Plus | 29128849-29128974 |
|    | 337489 | Dunham, I. et.al. | Plus | 33295724-33295872 |
|    | 337503 | Dunham, I. et.al. | Plus | 33385583-33385857 |
|    | 337504 | Dunham, I. et.al. | Plus | 33386053-33386236 |
|    | 337570 | Dunham, I. et.al. | Plus | 359309-359459     |
| 40 | 337585 | Dunham, I. et.al. | Plus | 951744-952008     |
|    | 337629 | Dunham, I. et.al. | Plus | 2017380-2017517   |
|    | 337670 | Dunham, I. et.al. | Plus | 3110593-3110760   |
|    | 337674 | Dunham, I. et.al. | Plus | 3332616-3332697   |
|    | 337692 | Dunham, I. et.al. | Plus | 3575105-3575299   |
| 45 | 337740 | Dunham, I. et.al. | Plus | 3870165-3870223   |
|    | 337755 | Dunham, I. et.al. | Plus | 3971764-3971900   |
|    | 337807 | Dunham, I. et.al. | Plus | 4444885-4444981   |
|    | 337844 | Dunham, I. et.al. | Plus | 4993372-4993603   |
|    | 337902 | Dunham, I. et.al. | Plus | 5682218-5682307   |
| 50 | 337904 | Dunham, I. et.al. | Plus | 5685819-5686012   |
|    | 337919 | Dunham, I. et.al. | Plus | 6035207-6035326   |
|    | 337951 | Dunham, I. et.al. | Plus | 6766321-6766382   |
|    | 337958 | Dunham, I. et.al. | Plus | 6969162-6969270   |
|    | 337964 | Dunham, I. et.al. | Plus | 7032720-7032802   |
| 55 | 338008 | Dunham, I. et.al. | Plus | 7697068-7697236   |
|    | 338053 | Dunham, I. et.al. | Plus | 8412742-8412823   |
|    | 338057 | Dunham, I. et.al. | Plus | 8526397-8526522   |
|    | 338059 | Dunham, I. et.al. | Plus | 8540638-8540712   |
|    | 338120 | Dunham, I. et.al. | Plus | 10765673-10765820 |
| 60 | 338124 | Dunham, I. et.al. | Plus | 10860311-10860471 |
|    | 338178 | Dunham, I. et.al. | Plus | 12800037-12800181 |
|    | 338196 | Dunham, I. et.al. | Plus | 13629317-13629466 |
|    | 338204 | Dunham, I. et.al. | Plus | 13870980-13871152 |
|    | 338239 | Dunham, I. et.al. | Plus | 14669918-14670016 |
| 65 | 338249 | Dunham, I. et.al. | Plus | 14870864-14870944 |
|    | 338250 | Dunham, I. et.al. | Plus | 14874504-14874575 |
|    | 338251 | Dunham, I. et.al. | Plus | 14963460-14963521 |
|    | 338260 | Dunham, I. et.al. | Plus | 15458919-15459257 |
|    | 338282 | Dunham, I. et.al. | Plus | 16240812-16241002 |
| 70 | 338316 | Dunham, I. et.al. | Plus | 17089711-17089988 |
|    | 338364 | Dunham, I. et.al. | Plus | 18210049-18210226 |
|    | 338374 | Dunham, I. et.al. | Plus | 18371200-18371282 |
|    | 338454 | Dunham, I. et.al. | Plus | 20180035-20180113 |
|    | 338494 | Dunham, I. et.al. | Plus | 21181818-21182009 |
| 75 | 338596 | Dunham, I. et.al. | Plus | 23078273-23078348 |
|    | 338622 | Dunham, I. et.al. | Plus | 23546552-23546749 |

|    |        |                   |       |                   |
|----|--------|-------------------|-------|-------------------|
|    | 338702 | Dunham, I. et.al. | Plus  | 25219632-25219739 |
|    | 338707 | Dunham, I. et.al. | Plus  | 25266346-25266417 |
|    | 338716 | Dunham, I. et.al. | Plus  | 25472519-25472686 |
| 5  | 338765 | Dunham, I. et.al. | Plus  | 26657278-26657346 |
|    | 338852 | Dunham, I. et.al. | Plus  | 28086911-28086971 |
|    | 338862 | Dunham, I. et.al. | Plus  | 28230332-28230444 |
|    | 338962 | Dunham, I. et.al. | Plus  | 29581892-29582020 |
|    | 338997 | Dunham, I. et.al. | Plus  | 30092658-30092730 |
| 10 | 339164 | Dunham, I. et.al. | Plus  | 32207441-32207802 |
|    | 339305 | Dunham, I. et.al. | Plus  | 33334676-33334864 |
|    | 339313 | Dunham, I. et.al. | Plus  | 33383457-33383585 |
|    | 339319 | Dunham, I. et.al. | Plus  | 33410900-33410972 |
|    | 339323 | Dunham, I. et.al. | Plus  | 33418663-33418829 |
|    | 339356 | Dunham, I. et.al. | Plus  | 33573387-33573517 |
| 15 | 339358 | Dunham, I. et.al. | Plus  | 33577760-33577922 |
|    | 339361 | Dunham, I. et.al. | Plus  | 33580121-33580251 |
|    | 339413 | Dunham, I. et.al. | Plus  | 34268734-34268875 |
|    | 339418 | Dunham, I. et.al. | Plus  | 34353362-34353421 |
|    | 339436 | Dunham, I. et.al. | Plus  | 34546469-34546834 |
| 20 | 332813 | Dunham, I. et.al. | Minus | 318840-318777     |
|    | 332854 | Dunham, I. et.al. | Minus | 1283611-1283053   |
|    | 332858 | Dunham, I. et.al. | Minus | 1339607-1339397   |
|    | 332863 | Dunham, I. et.al. | Minus | 1389980-1389884   |
|    | 332868 | Dunham, I. et.al. | Minus | 1413234-1413078   |
| 25 | 332884 | Dunham, I. et.al. | Minus | 1573063-1572923   |
|    | 332886 | Dunham, I. et.al. | Minus | 1574863-1574660   |
|    | 332896 | Dunham, I. et.al. | Minus | 1631641-1631422   |
|    | 332929 | Dunham, I. et.al. | Minus | 2020758-2020664   |
|    | 332930 | Dunham, I. et.al. | Minus | 2022565-2022497   |
| 30 | 332931 | Dunham, I. et.al. | Minus | 2023651-2023562   |
|    | 332932 | Dunham, I. et.al. | Minus | 2035348-2035282   |
|    | 332965 | Dunham, I. et.al. | Minus | 2537457-2537396   |
|    | 332995 | Dunham, I. et.al. | Minus | 2708847-2708685   |
|    | 333002 | Dunham, I. et.al. | Minus | 2537457-2537396   |
| 35 | 333011 | Dunham, I. et.al. | Minus | 2769669-2769571   |
|    | 333029 | Dunham, I. et.al. | Minus | 2885241-2885175   |
|    | 333033 | Dunham, I. et.al. | Minus | 2889900-2889699   |
|    | 333110 | Dunham, I. et.al. | Minus | 3244892-3244779   |
| 40 | 333126 | Dunham, I. et.al. | Minus | 3324305-3324184   |
|    | 333217 | Dunham, I. et.al. | Minus | 3967830-3967716   |
|    | 333220 | Dunham, I. et.al. | Minus | 3969363-3968789   |
|    | 333242 | Dunham, I. et.al. | Minus | 4104544-4104259   |
|    | 333243 | Dunham, I. et.al. | Minus | 4104961-4104728   |
| 45 | 333259 | Dunham, I. et.al. | Minus | 4306769-4306639   |
|    | 333270 | Dunham, I. et.al. | Minus | 4373573-4373219   |
|    | 333278 | Dunham, I. et.al. | Minus | 4414616-4414389   |
|    | 333279 | Dunham, I. et.al. | Minus | 4415252-4414844   |
|    | 333358 | Dunham, I. et.al. | Minus | 4732336-4732236   |
| 50 | 333408 | Dunham, I. et.al. | Minus | 4936879-4936661   |
|    | 333441 | Dunham, I. et.al. | Minus | 2708847-2708685   |
|    | 333444 | Dunham, I. et.al. | Minus | 5070077-5069643   |
|    | 333447 | Dunham, I. et.al. | Minus | 2537457-2537396   |
|    | 333466 | Dunham, I. et.al. | Minus | 2708847-2708685   |
| 55 | 333473 | Dunham, I. et.al. | Minus | 2537457-2537396   |
|    | 333496 | Dunham, I. et.al. | Minus | 5404643-5404523   |
|    | 333511 | Dunham, I. et.al. | Minus | 5557881-5557718   |
|    | 333542 | Dunham, I. et.al. | Minus | 5861529-5861341   |
|    | 333568 | Dunham, I. et.al. | Minus | 5965072-5964999   |
| 60 | 333582 | Dunham, I. et.al. | Minus | 6158522-6158322   |
|    | 333665 | Dunham, I. et.al. | Minus | 6975471-6975215   |
|    | 333673 | Dunham, I. et.al. | Minus | 7054704-7054602   |
|    | 333734 | Dunham, I. et.al. | Minus | 7535394-7535309   |
|    | 333964 | Dunham, I. et.al. | Minus | 8626045-8625966   |
| 65 | 334156 | Dunham, I. et.al. | Minus | 10580883-10580765 |
|    | 334172 | Dunham, I. et.al. | Minus | 11644142-11644008 |
|    | 334183 | Dunham, I. et.al. | Minus | 11832582-11832508 |
|    | 334184 | Dunham, I. et.al. | Minus | 11833848-11833757 |
|    | 334223 | Dunham, I. et.al. | Minus | 12734365-12734269 |
|    | 334270 | Dunham, I. et.al. | Minus | 13249131-13249007 |
| 70 | 334288 | Dunham, I. et.al. | Minus | 13295104-13294969 |
|    | 334303 | Dunham, I. et.al. | Minus | 13454331-13454217 |
|    | 334358 | Dunham, I. et.al. | Minus | 13724372-13724201 |
|    | 334370 | Dunham, I. et.al. | Minus | 13782655-13782493 |
|    | 334472 | Dunham, I. et.al. | Minus | 14391308-14391169 |
| 75 | 334474 | Dunham, I. et.al. | Minus | 14391920-14391809 |

|    |        |                   |       |                   |
|----|--------|-------------------|-------|-------------------|
|    | 334487 | Dunham, I. et.al. | Minus | 14432191-14432132 |
|    | 334500 | Dunham, I. et.al. | Minus | 14486730-14486621 |
|    | 334537 | Dunham, I. et.al. | Minus | 14827542-14827354 |
| 5  | 334621 | Dunham, I. et.al. | Minus | 15190418-15190299 |
|    | 334648 | Dunham, I. et.al. | Minus | 15363301-15363222 |
|    | 334764 | Dunham, I. et.al. | Minus | 16151208-16151104 |
|    | 334783 | Dunham, I. et.al. | Minus | 16293336-16293226 |
|    | 334784 | Dunham, I. et.al. | Minus | 16294548-16294360 |
| 10 | 334786 | Dunham, I. et.al. | Minus | 16297434-16297275 |
|    | 334789 | Dunham, I. et.al. | Minus | 16306095-16305996 |
|    | 334806 | Dunham, I. et.al. | Minus | 16433227-16433125 |
|    | 334823 | Dunham, I. et.al. | Minus | 16851360-16851189 |
|    | 334850 | Dunham, I. et.al. | Minus | 17660892-17660787 |
|    | 334924 | Dunham, I. et.al. | Minus | 19744615-19744229 |
| 15 | 334939 | Dunham, I. et.al. | Minus | 20131162-20131054 |
|    | 334943 | Dunham, I. et.al. | Minus | 20135064-20134903 |
|    | 334948 | Dunham, I. et.al. | Minus | 20141727-20141583 |
|    | 334970 | Dunham, I. et.al. | Minus | 20195886-20195554 |
|    | 334991 | Dunham, I. et.al. | Minus | 20341858-20341773 |
| 20 | 334993 | Dunham, I. et.al. | Minus | 20354277-20354174 |
|    | 335062 | Dunham, I. et.al. | Minus | 20921289-20921087 |
|    | 335207 | Dunham, I. et.al. | Minus | 21763011-21762880 |
|    | 335288 | Dunham, I. et.al. | Minus | 22304275-22303770 |
| 25 | 335289 | Dunham, I. et.al. | Minus | 22305950-22305708 |
|    | 335293 | Dunham, I. et.al. | Minus | 22316408-22316275 |
|    | 335332 | Dunham, I. et.al. | Minus | 22557778-22557557 |
|    | 335389 | Dunham, I. et.al. | Minus | 23043682-23043558 |
|    | 335478 | Dunham, I. et.al. | Minus | 23924778-23924329 |
| 30 | 335524 | Dunham, I. et.al. | Minus | 24237218-24236208 |
|    | 335547 | Dunham, I. et.al. | Minus | 24658526-24658460 |
|    | 335671 | Dunham, I. et.al. | Minus | 25358629-25358533 |
|    | 335682 | Dunham, I. et.al. | Minus | 25421215-25421093 |
|    | 335684 | Dunham, I. et.al. | Minus | 25425165-25425096 |
|    | 335698 | Dunham, I. et.al. | Minus | 25493029-25492767 |
| 35 | 335755 | Dunham, I. et.al. | Minus | 25763806-25763747 |
|    | 335756 | Dunham, I. et.al. | Minus | 25764330-25764251 |
|    | 335773 | Dunham, I. et.al. | Minus | 25880858-25880661 |
|    | 335813 | Dunham, I. et.al. | Minus | 26318734-26318649 |
| 40 | 335815 | Dunham, I. et.al. | Minus | 26320518-26320421 |
|    | 335817 | Dunham, I. et.al. | Minus | 26321875-26321750 |
|    | 335827 | Dunham, I. et.al. | Minus | 26380557-26380472 |
|    | 335829 | Dunham, I. et.al. | Minus | 26382348-26382251 |
|    | 335836 | Dunham, I. et.al. | Minus | 26397823-26397694 |
|    | 335857 | Dunham, I. et.al. | Minus | 26677208-26677096 |
| 45 | 335860 | Dunham, I. et.al. | Minus | 26684908-26684800 |
|    | 335871 | Dunham, I. et.al. | Minus | 26734972-26734892 |
|    | 335896 | Dunham, I. et.al. | Minus | 26977639-26977558 |
|    | 335897 | Dunham, I. et.al. | Minus | 26978293-26978142 |
| 50 | 335903 | Dunham, I. et.al. | Minus | 26985739-26985580 |
|    | 335914 | Dunham, I. et.al. | Minus | 27024197-27023994 |
|    | 335916 | Dunham, I. et.al. | Minus | 27027028-27026912 |
|    | 335935 | Dunham, I. et.al. | Minus | 27360288-27360058 |
|    | 335940 | Dunham, I. et.al. | Minus | 27420194-27420000 |
| 55 | 335999 | Dunham, I. et.al. | Minus | 28033986-28033848 |
|    | 336045 | Dunham, I. et.al. | Minus | 29044217-29044140 |
|    | 336140 | Dunham, I. et.al. | Minus | 30134204-30133980 |
|    | 336182 | Dunham, I. et.al. | Minus | 30371411-30371339 |
|    | 336198 | Dunham, I. et.al. | Minus | 30459668-30459460 |
| 60 | 336227 | Dunham, I. et.al. | Minus | 30902014-30901946 |
|    | 336246 | Dunham, I. et.al. | Minus | 31425669-31425253 |
|    | 336262 | Dunham, I. et.al. | Minus | 31833610-31833533 |
|    | 336292 | Dunham, I. et.al. | Minus | 32818035-32817927 |
|    | 336347 | Dunham, I. et.al. | Minus | 33843218-33843104 |
| 65 | 336397 | Dunham, I. et.al. | Minus | 34021504-34021389 |
|    | 336434 | Dunham, I. et.al. | Minus | 34073056-34072952 |
|    | 336662 | Dunham, I. et.al. | Minus | 2158060-2157993   |
|    | 336675 | Dunham, I. et.al. | Minus | 2020758-2020664   |
|    | 336676 | Dunham, I. et.al. | Minus | 2022565-2022497   |
| 70 | 336677 | Dunham, I. et.al. | Minus | 2023651-2023562   |
|    | 336678 | Dunham, I. et.al. | Minus | 2035348-2035282   |
|    | 336684 | Dunham, I. et.al. | Minus | 2158060-2157993   |
|    | 336703 | Dunham, I. et.al. | Minus | 5071373-5071278   |
|    | 336707 | Dunham, I. et.al. | Minus | 2820219-2820111   |
|    | 336722 | Dunham, I. et.al. | Minus | 3377722-3377590   |
| 75 | 336778 | Dunham, I. et.al. | Minus | 5071373-5071278   |

|    |        |                   |       |                   |
|----|--------|-------------------|-------|-------------------|
|    | 336846 | Dunham, I. et.al. | Minus | 7566306-7566238   |
|    | 336977 | Dunham, I. et.al. | Minus | 14110003-14109910 |
|    | 336981 | Dunham, I. et.al. | Minus | 14478638-14478472 |
| 5  | 337056 | Dunham, I. et.al. | Minus | 17975104-17974976 |
|    | 337094 | Dunham, I. et.al. | Minus | 20146915-20146778 |
|    | 337102 | Dunham, I. et.al. | Minus | 20581738-20581628 |
|    | 337161 | Dunham, I. et.al. | Minus | 23473450-23473375 |
|    | 337366 | Dunham, I. et.al. | Minus | 30961904-30961787 |
| 10 | 337422 | Dunham, I. et.al. | Minus | 32030671-32030417 |
|    | 337452 | Dunham, I. et.al. | Minus | 32415187-32415117 |
|    | 337534 | Dunham, I. et.al. | Minus | 34193847-34193769 |
|    | 337595 | Dunham, I. et.al. | Minus | 1020506-1020210   |
|    | 337602 | Dunham, I. et.al. | Minus | 1282987-1282741   |
| 15 | 337603 | Dunham, I. et.al. | Minus | 1299296-1299194   |
|    | 337645 | Dunham, I. et.al. | Minus | 5141462-5141329   |
|    | 337760 | Dunham, I. et.al. | Minus | 4008389-4008037   |
|    | 337772 | Dunham, I. et.al. | Minus | 4061918-4061782   |
|    | 337776 | Dunham, I. et.al. | Minus | 4084555-4084460   |
| 20 | 337840 | Dunham, I. et.al. | Minus | 4940540-4940409   |
|    | 337908 | Dunham, I. et.al. | Minus | 5697187-5697071   |
|    | 337937 | Dunham, I. et.al. | Minus | 6556005-6555907   |
|    | 337974 | Dunham, I. et.al. | Minus | 7153401-7153085   |
|    | 337986 | Dunham, I. et.al. | Minus | 7296008-7295951   |
| 25 | 338000 | Dunham, I. et.al. | Minus | 7530875-7530793   |
|    | 338221 | Dunham, I. et.al. | Minus | 14183649-14183568 |
|    | 338510 | Dunham, I. et.al. | Minus | 21339584-21339508 |
|    | 338535 | Dunham, I. et.al. | Minus | 21799696-21799274 |
|    | 338546 | Dunham, I. et.al. | Minus | 22012448-22012383 |
| 30 | 338556 | Dunham, I. et.al. | Minus | 22179326-22179234 |
|    | 338561 | Dunham, I. et.al. | Minus | 22311966-22311856 |
|    | 338668 | Dunham, I. et.al. | Minus | 24500606-24500442 |
|    | 338689 | Dunham, I. et.al. | Minus | 24893073-24892972 |
|    | 338727 | Dunham, I. et.al. | Minus | 25926788-25926580 |
| 35 | 338759 | Dunham, I. et.al. | Minus | 26582475-26582199 |
|    | 338763 | Dunham, I. et.al. | Minus | 26628148-26628009 |
|    | 338779 | Dunham, I. et.al. | Minus | 27030151-27029795 |
|    | 338876 | Dunham, I. et.al. | Minus | 28364326-28364071 |
|    | 338889 | Dunham, I. et.al. | Minus | 28477552-28477412 |
| 40 | 339028 | Dunham, I. et.al. | Minus | 30574122-30573937 |
|    | 339044 | Dunham, I. et.al. | Minus | 30721853-30721740 |
|    | 339128 | Dunham, I. et.al. | Minus | 31692815-31692686 |
|    | 339188 | Dunham, I. et.al. | Minus | 32347554-32347250 |
|    | 339208 | Dunham, I. et.al. | Minus | 32491714-32491657 |
| 45 | 339215 | Dunham, I. et.al. | Minus | 32502559-32502383 |
|    | 339230 | Dunham, I. et.al. | Minus | 32729004-32728929 |
|    | 339256 | Dunham, I. et.al. | Minus | 32926055-32925967 |
|    | 339280 | Dunham, I. et.al. | Minus | 33114230-33114010 |
|    | 339370 | Dunham, I. et.al. | Minus | 33805912-33805797 |
|    | 337895 |                   |       |                   |
| 50 | 329598 | 3962482           | Plus  | 39924-40220       |
|    | 329563 | 3962490           | Minus | 410-635           |
|    | 329557 | 3962492           | Minus | 53197-53647       |
|    | 329539 | 3983503           | Minus | 1-326             |
| 55 | 329526 | 3983506           | Plus  | 12251-12325       |
|    | 329524 | 3983507           | Minus | 38025-38143       |
|    | 329502 | 3983517           | Plus  | 75-338            |
|    | 329503 | 3983517           | Minus | 1801-1937         |
|    | 329499 | 3983518           | Plus  | 33463-33789       |
| 60 | 329479 | 3983526           | Minus | 7425-7561         |
|    | 329625 | 4567169           | Minus | 85893-85984       |
|    | 325363 | 5866920           | Plus  | 700446-700516     |
|    | 325366 | 5866920           | Minus | 920962-921713     |
|    | 325433 | 5866936           | Minus | 480706-480826     |
|    | 325447 | 5866941           | Minus | 372480-372621     |
| 65 | 325481 | 5866957           | Plus  | 47590-47672       |
|    | 325482 | 5866957           | Plus  | 47957-48078       |
|    | 325472 | 6017034           | Minus | 289581-289657     |
|    | 325513 | 6017035           | Minus | 34295-34490       |
| 70 | 325519 | 6017036           | Minus | 186804-186915     |
|    | 325587 | 6682462           | Plus  | 126724-126967     |
|    | 325585 | 6682462           | Plus  | 73476-73574       |
|    | 325594 | 5866992           | Minus | 470474-470566     |
|    | 325609 | 5866996           | Minus | 981751-981849     |
| 75 | 325622 | 5867000           | Plus  | 69994-70075       |
|    | 325751 | 6682474           | Plus  | 130437-130520     |

|    |        |         |       |                 |
|----|--------|---------|-------|-----------------|
|    | 325815 | 6682483 | Minus | 129273-130754   |
|    | 329762 | 6048280 | Plus  | 127744-127878   |
|    | 329789 | 6469354 | Minus | 118977-119036   |
| 5  | 329797 | 6523160 | Minus | 10616-10894     |
|    | 329838 | 6672062 | Plus  | 33990-34098     |
|    | 325864 | 5867069 | Minus | 110834-110904   |
|    | 325885 | 5867087 | Plus  | 193212-193377   |
|    | 325892 | 5867088 | Minus | 10498-10652     |
| 10 | 325929 | 5867125 | Minus | 51715-51996     |
|    | 325843 | 6552453 | Minus | 7126-7232       |
|    | 329989 | 4567166 | Plus  | 72861-73052     |
|    | 329960 | 5091594 | Minus | 1031-1162       |
|    | 329959 | 5103803 | Plus  | 188050-188193   |
|    | 329936 | 6165200 | Minus | 82761-82920     |
| 15 | 329919 | 6223624 | Minus | 103492-103681   |
|    | 330004 | 6623963 | Minus | 78872-78999     |
|    | 326025 | 5867176 | Plus  | 70854-70915     |
|    | 326054 | 5867184 | Minus | 146342-146469   |
| 20 | 326112 | 5867192 | Plus  | 2151-2725       |
|    | 326165 | 5867208 | Minus | 62787-62929     |
|    | 326213 | 5867224 | Minus | 60751-60927     |
|    | 326219 | 5867226 | Minus | 264008-264274   |
|    | 326160 | 5867254 | Minus | 112000-112137   |
| 25 | 326257 | 5867264 | Plus  | 222712-222819   |
|    | 330057 | 6478962 | Plus  | 75145-75287     |
|    | 326359 | 5867293 | Plus  | 9436-9494       |
|    | 326393 | 5867341 | Plus  | 41702-41841     |
|    | 326399 | 5867353 | Plus  | 6385-6536       |
| 30 | 326401 | 5867355 | Plus  | 35165-35332     |
|    | 326416 | 5867362 | Minus | 45283-45375     |
|    | 326431 | 5867371 | Plus  | 15855-15971     |
|    | 326460 | 5867400 | Minus | 142633-142935   |
|    | 326517 | 5867439 | Plus  | 44732-46356     |
| 35 | 326519 | 5867439 | Plus  | 166004-166243   |
|    | 326596 | 6138928 | Plus  | 133386-133563   |
|    | 330081 | 6015314 | Minus | 5768-5835       |
|    | 326714 | 5867595 | Plus  | 124490-124568   |
|    | 326752 | 5867615 | Minus | 1214-1562       |
| 40 | 326757 | 6249610 | Plus  | 74531-74597     |
|    | 326668 | 6552455 | Plus  | 146726-146838   |
|    | 326720 | 6552456 | Plus  | 84525-84677     |
|    | 326725 | 6552456 | Minus | 223005-223125   |
|    | 326862 | 6552465 | Plus  | 107702-107782   |
| 45 | 326882 | 6682509 | Minus | 167988-168179   |
|    | 326892 | 6682511 | Plus  | 119424-119500   |
|    | 326996 | 5867660 | Minus | 63212-63404     |
|    | 327010 | 5867664 | Plus  | 941057-941139   |
|    | 326919 | 6456782 | Minus | 40486-41046     |
| 50 | 327042 | 6531965 | Minus | 1380806-1381443 |
|    | 327049 | 6531965 | Minus | 1924026-1924110 |
|    | 327072 | 6531965 | Minus | 3796429-3797197 |
|    | 327074 | 6531965 | Plus  | 4039993-4040096 |
|    | 327075 | 6531965 | Plus  | 4041318-4041431 |
| 55 | 326981 | 6588016 | Plus  | 105091-106038   |
|    | 327133 | 6682522 | Plus  | 38069-38938     |
|    | 330137 | 4210430 | Minus | 21220-21377     |
|    | 330138 | 4210430 | Minus | 22334-22460     |
|    | 330143 | 4210430 | Plus  | 184737-184848   |
| 60 | 330153 | 4325335 | Plus  | 146951-147475   |
|    | 330135 | 4456470 | Minus | 121583-121885   |
|    | 327205 | 5867447 | Plus  | 167335-167576   |
|    | 327212 | 5867463 | Minus | 42308-42424     |
|    | 327287 | 5867479 | Minus | 62838-63024     |
| 65 | 327331 | 5867516 | Minus | 55606-55737     |
|    | 327364 | 6552412 | Minus | 115235-115396   |
|    | 327413 | 5867750 | Plus  | 101410-101508   |
|    | 327481 | 5867783 | Plus  | 104472-104673   |
|    | 327458 | 6004455 | Plus  | 173257-173378   |
| 70 | 327516 | 6117815 | Plus  | 199078-199216   |
|    | 327527 | 6381882 | Minus | 98950-99040     |
|    | 327548 | 5867797 | Minus | 81067-81130     |
|    | 327554 | 5867801 | Minus | 23092-23191     |
|    | 327565 | 5867811 | Plus  | 32516-32778     |
| 75 | 327600 | 6004462 | Minus | 2621-2862       |
|    | 327687 | 5867847 | Minus | 169293-169362   |

|    |        |         |       |                 |
|----|--------|---------|-------|-----------------|
|    | 330182 | 5123954 | Plus  | 120156-120245   |
|    | 327742 | 5867944 | Minus | 143307-143512   |
|    | 327805 | 5867968 | Plus  | 19952-20019     |
| 5  | 327809 | 5867968 | Plus  | 54610-54761     |
|    | 327814 | 5867968 | Plus  | 69377-70566     |
|    | 327815 | 5867968 | Plus  | 70804-71401     |
|    | 327791 | 5867977 | Plus  | 22491-22610     |
|    | 327745 | 6531959 | Minus | 229066-229124   |
| 10 | 330211 | 6013592 | Plus  | 59158-59215     |
|    | 330207 | 6013606 | Minus | 109912-110004   |
|    | 330257 | 6671881 | Minus | 143228-143393   |
|    | 330262 | 6671884 | Plus  | 67913-68053     |
|    | 330286 | 6671913 | Minus | 31050-31171     |
| 15 | 328105 | 5868020 | Minus | 301705-301784   |
|    | 328113 | 5868024 | Minus | 80378-80491     |
|    | 328142 | 5868050 | Minus | 9656-9778       |
|    | 328152 | 5868060 | Minus | 73981-74203     |
|    | 328170 | 5868071 | Plus  | 93170-93295     |
|    | 327910 | 5868162 | Plus  | 21622-21748     |
| 20 | 327919 | 5868165 | Plus  | 547701-547800   |
|    | 327990 | 5868218 | Minus | 36225-36503     |
|    | 328249 | 6381891 | Minus | 96352-96527     |
|    | 328251 | 6381891 | Plus  | 124444-124557   |
| 25 | 328253 | 6381894 | Minus | 4411-4509       |
|    | 328084 | 6469819 | Minus | 155366-155459   |
|    | 328274 | 5868219 | Minus | 31244-31439     |
|    | 328615 | 5868239 | Plus  | 35214-35347     |
|    | 328632 | 5868247 | Plus  | 76734-76853     |
| 30 | 328779 | 5868309 | Plus  | 41570-41639     |
|    | 328783 | 5868309 | Minus | 73658-73822     |
|    | 328801 | 5868321 | Minus | 44492-44609     |
|    | 328820 | 5868330 | Plus  | 90446-90602     |
|    | 328835 | 5868339 | Plus  | 88053-88461     |
| 35 | 328290 | 5868363 | Minus | 127366-127496   |
|    | 328321 | 5868373 | Minus | 1029614-1029673 |
|    | 328332 | 5868375 | Plus  | 280154-280289   |
|    | 328333 | 5868375 | Plus  | 282506-282664   |
|    | 328349 | 5868383 | Minus | 260704-260804   |
| 40 | 328450 | 5868425 | Minus | 209192-209321   |
|    | 328466 | 5868434 | Minus | 15643-15900     |
|    | 328479 | 5868449 | Minus | 331-560         |
|    | 328481 | 5868449 | Minus | 8987-9180       |
|    | 328546 | 5868487 | Minus | 17547-17722     |
| 45 | 328662 | 6004473 | Plus  | 1184773-1184855 |
|    | 328767 | 6017031 | Minus | 35625-35723     |
|    | 328768 | 6017031 | Minus | 223741-224238   |
|    | 328857 | 6381927 | Minus | 80557-81051     |
|    | 328878 | 6552423 | Plus  | 105580-105774   |
| 50 | 328882 | 6552423 | Minus | 157669-157826   |
|    | 328690 | 6588001 | Minus | 571207-571274   |
|    | 328691 | 6588001 | Minus | 579598-579664   |
|    | 330307 | 4877982 | Plus  | 107384-107559   |
|    | 328903 | 5868514 | Plus  | 23625-24468     |
| 55 | 328987 | 5868535 | Minus | 25705-25764     |
|    | 328998 | 5868538 | Plus  | 40996-41104     |
|    | 329062 | 5868590 | Minus | 58977-59094     |
|    | 329086 | 5868604 | Minus | 35489-35588     |
|    | 329154 | 5868686 | Minus | 200851-201356   |
| 60 | 329156 | 5868686 | Minus | 202013-202341   |
|    | 329164 | 5868691 | Plus  | 62305-62517     |
|    | 329170 | 5868693 | Plus  | 67924-68019     |
|    | 329179 | 5868704 | Plus  | 181639-181815   |
|    | 329193 | 5868716 | Plus  | 168095-168181   |
| 65 | 329254 | 5868733 | Plus  | 4133-4214       |
|    | 329369 | 5868842 | Minus | 121148-121516   |
|    | 329367 | 5868842 | Minus | 87201-87587     |
|    | 329141 | 6017060 | Plus  | 343924-343997   |
|    | 329347 | 6456785 | Plus  | 18433-18897     |
| 70 | 329017 | 6682532 | Minus | 255591-255672   |
|    | 329434 | 5868883 | Minus | 31124-31263     |

**TABLE 2 DNA AND PROTEIN SEQUENCES FOR CBF9 AND BFO8**

Table 2 provides the nucleic acid and protein sequence of the CBF9 and BFO8 genes as well as the Unigene and Exemplar accession numbers for CBF9 and BFO8.

**CBF9 DNA SEQUENCE**

Gene name: ESTs  
 Unigene number: Hs.157601  
 Probeset Accession #: W07459  
 Nucleic Acid Accession #: AC005383  
 Coding Sequence: 328-2751 (underlined sequences correspond to start and stop codons)

|    |            |            |            |            |            |             |      |
|----|------------|------------|------------|------------|------------|-------------|------|
| 15 | 1          | 11         | 21         | 31         | 41         | 51          |      |
|    |            |            |            |            |            |             |      |
|    | GACAGTGTTC | GCGGCTGCAC | CGCTCGGAGG | CTGGGTGACC | CGCGTAGAAG | TGAAGTACTT  | 60   |
|    | TTTATATTGC | AGACCTGGGC | CGATGCCGCT | TTAAAAAAGC | CGAGGGGCTC | TATGCACCTC  | 120  |
|    | CCTGGCGGTA | GTTCTCTCGA | CCTCAGCCGG | GTGCGGTCTG | GCCGCCCTCT | CCCAGGAGAG  | 180  |
| 20 | ACAAACAGGT | GTCCACCGTG | GCAGCCGCGC | CCCGGGCGCC | CCTCCTGTGA | TCCCGTAGCG  | 240  |
|    | CCCCCTGGCC | CGAGCCGCGC | CCGGGTCTGT | GAGTAGAGCC | GCCCGGGCAC | CGAGCGCTGG  | 300  |
|    | TCGCCGCTCT | CCTTCCGTTA | TATCAACATG | CCCCCTTTC  | TGTTGCTGGA | GGCCGTCTGT  | 360  |
|    | GTTTCTCTGT | TTTCCAGAGT | GCCCCATCT  | CTCCCTCTCC | AGGAAGTCCA | TGTAAGCAAA  | 420  |
|    | GAAACCATCG | GGAAGATTTC | AGCTGCCAGC | AAAATGATGT | GGTGCTCGGC | TGCAGTGGAC  | 480  |
| 25 | ATCATGTTTC | TGTTAGATGG | GTCTAACAGC | GTGCGGAAAG | GGAGCTTTGA | AAGGTCCAAG  | 540  |
|    | CACTTTGCCA | TCACAGTCTG | TGACGGTCTG | GACATCAGCC | CCGAGAGGGT | CAGAGTGGGA  | 600  |
|    | GCAATCCAGT | TCAGTTCAC  | TCCTCATCTG | GAATTCCTT  | TGGATTCATT | TTCAACCCAA  | 660  |
|    | CAGAAAGTGA | AGGCAAGAAT | CAAGAGGATG | GATTTCAAAG | GAGGGCGCAC | GGAGACGGAA  | 720  |
|    | CTTGCTCTGA | AATACCTTCT | GCACAGAGGG | TTGCCTGGAG | GCAGAAATGC | TTCTGTGCCC  | 780  |
| 30 | CAGATCCTCA | TCATCCTCAC | TGATGGGAAG | TCCCAGGGGG | ATGTGGCACT | GCCATCCAAG  | 840  |
|    | CAGCTGAAGG | AAAGGGGTGT | CACTGTGTTT | GCTGTGGGGG | TCAGGTTTTC | CAGGTGGGAG  | 900  |
|    | GAGCTGCATG | CACTGGCCAG | CGAGCCTAGA | GGGCAGCACG | TGCTGTTGGC | TGAGCAGGTG  | 960  |
|    | GAGGATGCCA | CCAACGGCCT | CTTCAGCACC | CTCAGCAGCT | CGCCCATCTG | CTCCAGCGCC  | 1020 |
|    | ACGCCAGACT | GCAGGGTCTG | GGCTCACCCC | TGTGAGCACA | GGACGCTGGA | GATGGTCCGG  | 1080 |
| 35 | GAGTTCGCTG | GCAATGCCCC | ATGCTGGAGA | GGATCGCGGC | GGACCTTTCG | GGTGCTGGCT  | 1140 |
|    | GCACACTGTC | CCTTCTACAG | CTGGAAGAGA | GTGTTCTTAA | CCCACCCTGC | CACCTGCTAC  | 1200 |
|    | AGGACCACCT | GCCCAGGGCC | CTGTGACTCG | CAGCCCTGCC | AGAATGGAGG | CACATGTGTT  | 1260 |
|    | CCAGAAGGAC | TCCGTGGTGG | CCAGTGCCCT | TGCCCGCTGG | CCTTTGGAGG | GGAGGCTAAC  | 1320 |
|    | TGTGCCCTGA | AGCTGAGCCT | GGAATGCAGG | GTCGACCTCC | TCTTCCTGCT | GGACAGCTCT  | 1380 |
| 40 | GCGGGACCA  | CTCTGGACGG | CTTCTGCGG  | GCCAAAGTCT | TCGTGAAGCG | GTTTGTGCGG  | 1440 |
|    | GCCGTGCTGA | GCGAGGACTC | TGGGGCCCGA | GTGGGTGTGG | CCACATACAG | CAGGGAGCTG  | 1500 |
|    | CTGGTGGCGG | TGCGTGTGGG | GGAGTACCAG | GATGTGCCTG | ACCTGGTCTG | GAGCCTCGAT  | 1560 |
|    | GGCATTCCCT | TCCGTGGTGG | CCCCACCCTG | ACGGGCAGTG | CCTTGCGGCA | GGCGGCAGAG  | 1620 |
|    | CGTGGCTTCG | GGAGCGCCAC | CAGGACAGGC | CAGGACCGGC | CACGTAGAGT | GGTGGTTTTG  | 1680 |
| 45 | CTCACTGAGT | CACACTCCGA | GGATGAGGTT | GCGGGCCAG  | CGCGTCACGC | AAGGGCGCGA  | 1740 |
|    | GAGCTGCTCC | TGCTGGGTGT | AGGCAGTGAG | GCCGTGCGGG | CAGAGCTGGA | GGAGATCACA  | 1800 |
|    | GGCAGCCCAA | AGCATGTGAT | GGTCTACTCG | GATCCTCAGG | ATCTGTTCAA | CCAAATCCCT  | 1860 |
|    | GAGCTGCAGG | GGAAGCTGTG | CAGCCGGCAG | CGGCCAGGGT | GCCGGACACA | AGCCCTGGAC  | 1920 |
|    | CTCGTCTTCA | TGTTGGACAC | CTCTGCCTCA | GTAGGGCCCG | AGAATTTTGC | TCAGATGCAG  | 1980 |
| 50 | AGCTTTGTGA | GAAGCTGTGC | CCTCCAGTTT | GAGGTGAACC | CTGACGTGAC | ACAGGTGCGC  | 2040 |
|    | CTGGTGGTGT | ATGGCAGCCA | GGTGCAGACT | GCCTTCGGGC | TGACACCAA  | ACCCACCCGG  | 2100 |
|    | GCTGCGATGC | TGCGGGCCAT | TAGCCAGGCG | CCCTACCTAG | GTGGGGTGGG | CTCAGCCGGC  | 2160 |
|    | ACGCCCTGCG | TGCACATCTA | TGACAAAGTG | ATGACCGTCC | AGAGGGGTGC | CCGGCCTGCT  | 2220 |
|    | GTCCCAAAAG | CTGTGGTGGT | GCTCACAGGC | GGGAGAGGCG | CAGAGGATGC | AGCCGTTCTT  | 2280 |
| 55 | GCCAGAAGC  | TGAGGAACAA | TGGCATCTCT | GTCTTGGTCG | TGGGCGTGGG | GCCTGTCCTA  | 2340 |
|    | AGTGAGGGTC | TGCGGAGGCT | TGCAGGTCCC | CGGGATTCCC | TGATCCACGT | GGCAGCTTAC  | 2400 |
|    | GCCGACCTGC | GGTACCACCA | GGACGTGCTC | ATTGAGTGGC | TGTGTGGAGA | AGCCAAGCAG  | 2460 |
|    | CCAGTCAACC | TCTGCAAAAC | CAGCCCGTGC | ATGAATGAGG | GCAGCTGCGT | CCTGCAGAAAT | 2520 |
|    | GGGAGCTACC | GCTGCAAGTG | TCGGGATGGC | TGGGAGGGCC | CCCACTGCGA | GAACCGTGAG  | 2580 |
| 60 | TGGAGCTCTT | GCTCTGTATG | TGTGAGCCAG | GGATGGATTG | TTGAGACGCC | CCTGAGGCAC  | 2640 |
|    | ATGCTCTCCG | TGCAGGAGGG | CAGCAGCCGT | ACCCCTCCCA | GCAACTACAG | AGAAGGCCTG  | 2700 |
|    | GGCACTGAAA | TGGTGCTTAC | CTTCTGGAAT | GTCTGTGCCC | CAGGTCTCTA | GAATGTCTGC  | 2760 |
|    | TTCCCCCGGT | GGCCAGGACC | ACTATTCTCA | CTGAGGGAGG | AGGATGTCCC | AACTGCAGCC  | 2820 |

ATGCTGCTTA GAGACAAGAA AGCAGCTGAT GTCACCCACA AACGATGTTG TTGAAAAGTT 2880  
 TTGATGTGTA AGTAAATACC CACTTTCCTGT ACCTGCTGTG CCTTGTGAG GCTATGTCAT 2940  
 CTGCCACCTT TCCCTTGAGG ATAAACAAGG GGTCTGAAG ACTTAAATTT AGCGGCCTGA 3000  
 CGTTCCTTTG CACACAATCA ATGCTCGCCA GAATGTTGTT GACACAGTAA TGCCCAGCAG 3060  
 5 AGGCCTTTAC TAGAGCATCC TTTGGACGGC GAAGGCCACG GCCTTTCAAG ATGGAAAGCA 3120  
 GCAGCTTTTC CACTTCCCCA GAGACATTCT GGATGCATTT GCATTGAGTC TGAAAGGGGG 3180  
 CTTGAGGGAC GTTTGTGACT TCTTGGCGAC TGCCTTTTGT GTGTGGAAGA GACTTGAAAA 3240  
 GGTCTCAGAC TGAATGTGAC CAATTAACCA GCTTGGTTGA TGATGGGGGA GGGGCTGAGT 3300  
 10 TGTGCATGGG CCCAGGTCTG GAGGGCCACG TAAAATCGTT CTGAGTCGTG AGCAGTGTCC 3360  
 ACCTTGAAGG TCTTC

### CBF9 Protein sequence

Gene name:

ESTs

Unigene number:

Hs.157601

Protein Accession #: none found

Signal sequence:

1-17

Transmembrane domains:

none found

VGW domains:

49-223; 341-518; 529-706

EGF domains:

298-333; 715-748

Cellular Localization:

plasma membrane

1 11 21 31 41 51  
 | | | | |  
 25 MPPFLLLEAV CVFLFSRVPP SLPLQEVHVS KETIGKISAA SKMMWCSAAV DIMFLLDGSN 60  
 SVGKGSFERS KHFAITVCDG LDISPERVRV GAFQFSSTPH LEFFLDSFST QQEVKARIKR 120  
 MVFKGGRRET ELALKYLLHR GLPGGRNASV PQILIIIVTDG KSQGDVALPS KQLKERGVTV 180  
 FAVGVRFPRW EELHALASEP RGQHVLLAEQ VEDATNGLFS TLSSSAICSS ATPDCRVEAH 240  
 PCEHRTLEMV REFAGNAPCW RGSRRTLAVL AAHCPFYSWK RVFLTHPATC YRTTCPGPCD 300  
 30 SQPCQNGGTC VPEGLDGYQC LCPLAFGGEA NCALKLSLEC RVDLLFLDLS SAGTTLDGFL 360  
 RAKVVFVKRFV RAVLSEDSRA RVGVATYSRE LLVAVPVGEY QDVPDLVWSL DGIPFRGGPT 420  
 LTGSALRQAA ERFGFSATRT QQDRPRRVVV LLTESHSEDE VAGPARHARA RELLLLGVGS 480  
 EAVRAELEEI TGSPKHMVY SDPQDLFNQI PELQGKLCR QRPGCRTQAL DLVFMMLDTSA 540  
 SVGPENFAQM QSFVRSCALQ FEVNPVDVTQV GLVVYGSQVQ TAFGLDTKPT RAAMLRAISQ 600  
 35 APYLGGVGSA GTALLHIYDK VMTVQRGARP GVPKAVVLT GGRGAEDAAV PAQKLNRNGI 660  
 SVLVVGVGPV LSEGLRRLAG PRDSLHVA YADLRYHQDV LIEWLCGEAK QPVNLCKPSP 720  
 CMNEGSCVLQ NGSYRCKCRD GWEGPHCENR EWSSCSVCVS QGWILETPLR HMAPVQEGSS 780  
 RTPPSNYREG LGTEMVPTFW NVCAPGP

### BFO8 DNA SEQUENCE

Gene name:

TMPRSS3a

Unigene number:

Hs.298241

Probeset Accession #:

AI538613

Nucleic Acid Accession #:

AB038157

Coding sequence:

202-1566 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 | | | | |  
 55 ACCGGGCACC GGACGGCTCG GGTACTTTCG TTCTTAATTA GGTCAATGCC GTGTGAGCCA 60  
 GGAAAGGGCT GTGTTTATGG GAAGCCAGTA ACACGTGGGC TACTATCTC TTCCGTGGTG 120  
 CCATCTACAT TTTTGGGACT CGGGAATTAT GAGGTAGAGG TGGAGGCGGA GCCGGATGTC 180  
 AGAGGTCTCTG AAATAGTCAC CATGGGGGAA AATGATCCGC CTGCTGTTGA AGCCCCCTTC 240  
 TCATTCCGAT CGCTTTTGG CTTTGATGAT TTGAAAATAA GTCTGTGTC ACCAGATGCA 300  
 60 GATGCTGTTG CTGCACAGAT CCTGTCACTG CTGCCATTGA AGTTTTTTCC AATCATCGTC 360  
 ATTGGGATCA TTGCATTGAT ATTAGCACTG GCCATTGGTC TGGGCATCCA CTTCGACTGC 420  
 TCAGGGAAGT ACAGATGTCG CTCATCCTTT AAGTGTATCG AGCTGATAGC TCGATGTGAC 480  
 GGAGTCTCGG ATTGCAAAGA CGGGGAGGAC GAGTACCGCT GTGTCCGGGT GGGTGGTCAG 540  
 AATGCCGTGC TCCAGGTGTT CACAGCTGCT TCGTGGAAGA CCATGTGCTC CGATGACTGG 600  
 AAGGGTCACT ACGCAAATGT TGCCTGTGCC CAACTGGGTT TCCCAAGCTA TGTGAGTTCA 660  
 65 GATAACCTCA GAGTGAGCTC GCTGGAGGG CAGTTCGGG AGGAGTTTGT GTCCATCGAT 720  
 CACCTCTTGC CAGATGACAA GGTGACTGCA TTACACCACT CAGTATATGT GAGGGAGGGA 780

|    |            |            |            |             |             |             |      |
|----|------------|------------|------------|-------------|-------------|-------------|------|
|    | TGTGCCTCTG | GCCACGTGGT | TACCTTGCAG | TGCACAGCCT  | GTGGTCATAG  | AAGGGGCTAC  | 840  |
|    | AGCTCACGCA | TCGTGGGTGG | AAACATGTCC | TTGCTCTCGC  | AGTGGCCCTG  | GCAGGCCAGC  | 900  |
|    | CTTCAGTTCC | AGGGCTACCA | CCTGTGCGGG | GGCTCTGTCA  | TCACGCCCCCT | GTGGATCATC  | 960  |
| 5  | ACTGCTGCAC | ACTGTGTTTA | TGACTTGTAC | CTCCCCAAGT  | CATGGACCAT  | CCAGGTGGGT  | 1020 |
|    | CTAGTTTCCC | TGTTGGACAA | TCCAGCCCCA | TCCCACCTGG  | TGGAGAAGAT  | TGTCTACCAC  | 1080 |
|    | AGCAAGTACA | AGCCAAAGAG | GCTGGGCAAT | GACATCGCCC  | TTATGAAGCT  | GGCCGGGCCA  | 1140 |
|    | CTCACGTTCA | ATGAAATGAT | CCAGCCTGTG | TGCCTGCCCA  | ACTCTGAAGA  | GAACCTCCCC  | 1200 |
|    | GATGGAAAAG | TGTGCTGGAC | GTCAGGATGG | GGGGCCACAG  | AGGATGGAGC  | AGGTGACGCC  | 1260 |
| 10 | TCCCCTGTCC | TGAACCACGC | GGCCGTCCCT | TTGATTTCCA  | ACAAGATCTG  | CAACCACAGG  | 1320 |
|    | GACGTGTACG | GTGCGATCAT | CTCCCCCTCC | ATGCTCTGCG  | CGGGCTACCT  | GACGGGTGGC  | 1380 |
|    | GTGGACAGCT | GCCAGGGGGA | CAGCGGGGGG | CCCCTGGTGT  | GTCAAGAGAG  | GAGGCTGTGG  | 1440 |
|    | AAGTTAGTGG | GAGCGACCAG | CTTTGGCATC | GGCTGCGCAG  | AGGTGAACAA  | GCCTGGGGTG  | 1500 |
|    | TACACCCGTG | TCACCTCCTT | CCTGGACTGG | ATCCACGAGC  | AGATGGAGAG  | AGACCTAAAA  | 1560 |
|    | ACCTGAAGAG | GAAGGGGACA | AGTAGCCACC | TGAGTTCTCTG | AGGTGATGAA  | GACAGCCCGA  | 1620 |
| 15 | TCTTCCCCTG | GACTCCCCTG | TAGGAACCTG | CACACGAGCA  | GACACCCTTG  | GAGCTCTGAG  | 1680 |
|    | TTCCGGCACC | AGTAGCAGGC | CCGAAAGAGG | CACCCCTTCCA | TCTGATTCCA  | GCACAACCTT  | 1740 |
|    | CAAGCTGCTT | TTTGTTTTTT | GTTTTTTTGA | GGTGGAGTCT  | CGCTCTGTTG  | CCCAGGCTGG  | 1800 |
|    | AGTGCAGTGG | CGAAATCCCT | GCTCACTGCA | GCCTCCGCTT  | CCCTGGTTCA  | AGCGATTCTC  | 1860 |
| 20 | TTGCCTCAGC | TTCCCCAGTA | GCTGGGACCA | CAGGTGCCCG  | CCACCACACC  | CAACTAATTT  | 1920 |
|    | TTGTATTTTT | AGTAGAGACA | GGGTTTCACC | ATGTTGGCCA  | GGCTGCTCTC  | AAACCCCTGA  | 1980 |
|    | CCTCAAATGA | TGTGCCTGCT | TCAGCCTCCC | ACAGTGCTGG  | GATTACAGGC  | ATGGGCCACC  | 2040 |
|    | ACGCCTAGCC | TCACGCTCCT | TTCTGATCTT | CACATAAGAAC | AAAAGAAGCA  | GCAACTTGCA  | 2100 |
|    | AGGGCGGCCT | TTCCCCTGCT | TCCATCTGGT | TTTCTCTCCA  | GGGGTCTTGC  | AAAATTCCCTG | 2160 |
| 25 | ACGAGATAAG | CAGTTATGTG | ACCTCACGTG | CAAAGCCACC  | AACAGCCACT  | CAGAAAAGAC  | 2220 |
|    | GCACCAGCCC | AGAAGTGCAG | AACTGCAGTC | ACTGCACGTT  | TTCATCTCTA  | GGGACCAGAA  | 2280 |
|    | CCAAACCCAC | CCTTTCTACT | TCCAAGACTT | ATTTTCACAT  | GTGGGGAGGT  | TAATCTAGGA  | 2340 |
|    | ATGACTCGTT | TAAAGCCTAT | TTTCATGATT | TCTTTGTAGC  | ATTTGGTGCT  | TGACGTATTA  | 2400 |
|    | TTGTCCTTTG | ATTCCAAATA | ATATGTTTCC | TTCCCTCAAA  | AAAAAAAAAA  | AAAAAAAAAA  | 2460 |
| 30 | AAAAAAA    |            |            |             |             |             |      |

**BFO8 Protein sequence:**

|    |                        |                |
|----|------------------------|----------------|
| 35 | Gene name:             | TMPRSS3a       |
|    | Unigene number:        | Hs.298241      |
|    | Probeset Accession #:  | AI538613       |
|    | Protein Accession #:   | BAB20077       |
|    | Signal sequence:       | none found     |
| 40 | Transmembrane domains: | 43-65, 239-261 |
|    | Tryp_Spc domain:       | 216-444        |
|    | Cellular Localization: | not determined |

|    |            |            |            |             |            |            |     |
|----|------------|------------|------------|-------------|------------|------------|-----|
| 45 | 1          | 11         | 21         | 31          | 41         | 51         |     |
|    |            |            |            |             |            |            |     |
|    | MGENDPPAVE | APFSFRSLFG | LDDLKISPVA | PDADAVAAQI  | LSLLPLKFFP | IIVIGIIALI | 60  |
|    | LALAIGLGIH | FDCSGKYRCR | SSFKCIELIA | RCDGVS DCKD | GEDEYRCVRV | GGQNAVLQVF | 120 |
|    | TAASWKTMC  | DDWKGHYANV | ACAQLGFPSY | VSSDNLRVSS  | LEGQFREEFV | SIDHLLPDDK | 180 |
| 50 | VTALHHSVYV | REGCASGHV  | TLQCTACGHR | RGYSSRIVGG  | NMSLLSQWPW | QASLQFQGYH | 240 |
|    | LCGGSVITPL | WIITAAHCYV | DLYLPKSWTI | QVGLVSLLDN  | PAPSHLVEKI | VYHSKYKPKR | 300 |
|    | LGNDIALMKL | AGPLTFNEMI | QPVCLPNSEE | NFPDGKVCWT  | SGWGATEDGA | GDASPVLNHA | 360 |
|    | AVPLISNKIC | NHRDVYGGII | SPSMLCAGYL | TGGVDSQCGD  | SGGPLVCQER | RLWKLVGATS | 420 |
| 55 | FGIGCAEVNK | PGVYTRVTSF | LDWIHEQMER | DLKT        |            |            |     |

WHAT IS CLAIMED IS:

- 1                   1.       A method of screening drug candidates comprising:
  - 2                   a) providing a cell that expresses an expression profile gene selected from the
  - 3                   group consisting of an expression profile gene set forth in Table 1 or Table 2 or fragment
  - 4                   thereof;
  - 5                   b) adding a drug candidate to said cell; and
  - 6                   c) determining the effect of said drug candidate on the expression of said
  - 7                   expression profile gene.
- 1                   2.       A method according to claim 1 wherein said determining comprises
- 2                   comparing the level of expression in the absence of said drug candidate to the level of
- 3                   expression in the presence of said drug candidate.
- 1                   3.       A method of screening for a bioactive agent capable of binding to a
- 2                   colorectal cancer modulator protein (colorectal cancer modulator protein), wherein said
- 3                   colorectal cancer modulator protein is encoded by a nucleic acid selected from the group
- 4                   consisting of a nucleic acid of Table 1 or Table 2 or a fragment thereof, said method
- 5                   comprising:
  - 6                   a) combining said colorectal cancer modulator protein and a candidate
  - 7                   bioactive agent; and
  - 8                   b) determining the binding of said candidate agent to said colorectal cancer
  - 9                   modulator protein.
- 1                   4.       A method for screening for a bioactive agent capable of modulating the
- 2                   activity of a colorectal cancer modulator protein, wherein said colorectal cancer modulator
- 3                   protein is encoded by a nucleic acid selected from the group consisting of a nucleic acid of
- 4                   Table 1 or Table 2 or a fragment thereof, said method comprising:
  - 5                   a) combining said colorectal cancer modulator protein and a candidate
  - 6                   bioactive agent; and

b) determining the effect of said candidate agent on the bioactivity of said colorectal cancer modulator protein.

5. A method of evaluating the effect of a candidate colorectal cancer drug comprising:

a) administering said drug to a patient;

b) removing a cell sample from said patient; and

c) determining the expression of a gene selected from the group consisting of a nucleic acid of Table 1 or Table 2.

6. A method according to claim 5 further comprising comparing said expression profile to an expression profile of a healthy individual.

7. A method of diagnosing colorectal cancer comprising:

a) determining the expression of one or more genes selected from the group consisting of a nucleic acid of Table 1 or Table 2 or a fragment thereof or a polypeptide encoded thereby in a first tissue type of a first individual; and

b) comparing said expression of said gene(s) from a second normal tissue type from said first individual or a second unaffected individual;

wherein a difference in said expression indicates that the first individual has colorectal cancer.

8. A method for screening for a bioactive agent capable of interfering with the binding of a colorectal cancer modulator protein (colorectal cancer modulator protein) or a fragment thereof and an antibody which binds to said colorectal cancer modulator protein or fragment thereof, said method comprising:

a) combining a colorectal cancer modulator protein or fragment thereof, a candidate bioactive agent and an antibody which binds to said colorectal cancer modulator protein or fragment thereof; and

b) determining the binding of said colorectal cancer modulator protein or fragment thereof and said antibody.

1                   9. A method for inhibiting the activity of a colorectal cancer modulator  
2 protein (colorectal cancer modulator protein), wherein said colorectal cancer modulator  
3 protein is encoded by a nucleic acid selected from the group consisting of a nucleic acid of  
4 Table 1 or Table 2 or a fragment thereof, said method comprising binding an inhibitor to said  
5 colorectal cancer modulator protein.

1                   10. A method according to claim 9 wherein said inhibitor is an antibody.

1                   11. A method of treating colorectal cancer comprising administering to a  
2 patient an inhibitor of a colorectal cancer modulator protein, wherein said colorectal cancer  
3 modulator protein is encoded by a nucleic acid selected from the group consisting of a  
4 nucleic acid of Table 1 or Table 2 or a fragment thereof.

1                   12. A method according to claim 11 wherein said inhibitor is an antibody.

1                   13. A method of neutralizing the effect of a colorectal cancer modulator  
2 protein, or a fragment thereof, comprising contacting an agent specific for said protein with  
3 said protein in an amount sufficient to effect neutralization.

1                   14. A method for localizing a therapeutic moiety to colorectal cancer tissue  
2 comprising exposing said tissue to an antibody to a colorectal cancer modulator protein or  
3 fragment thereof conjugated to said therapeutic moiety.

1                   15. The method of Claim 14, wherein said therapeutic moiety is a cytotoxic  
2 agent.

1                   16. The method of Claim 14, wherein said therapeutic moiety is a  
2 radioisotope.

1                   17. A method for inhibiting colorectal cancer in a cell, wherein said method  
2 comprises administering to a cell a composition comprising antisense molecules to a nucleic  
3 acid of Table 1 or Table 2.

1                   18. An antibody which specifically binds to a protein encoded by a nucleic  
2 acid of Table 1 or Table 2 or a fragment thereof.

1 19. The antibody of Claim 18, wherein said antibody is a monoclonal  
2 antibody.

1 20. The antibody of Claim 18, wherein said antibody is a humanized  
2 antibody.

1 21. The antibody of Claim 18, wherein said antibody is an antibody fragment.

1 22. A biochip comprising one or more nucleic acid segments selected from  
2 the group consisting of a nucleic acid of Table 1 or Table 2 or a fragment thereof, wherein  
3 said biochip comprises fewer than 1000 nucleic acid probes.

1 23. A nucleic acid having a sequence at least 95% homologous to a sequence  
2 of a nucleic acid of Table 1 or Table 2 or its complement.

1 24. A nucleic acid which hybridizes under high stringency to a nucleic acid of  
2 Table 1 or Table 2 or its complement.

1 25. A polypeptide encoded by the nucleic acid of Claim 23 or 24.

1 26. A method of eliciting an immune response in an individual, said method  
2 comprising administering to said individual a composition comprising the polypeptide of  
3 Claim 25 or a fragment thereof.

1 27. A method of eliciting an immune response in an individual, said method  
2 comprising administering to said individual a composition comprising a nucleic acid  
3 comprising a sequence of a nucleic acid of Table 1 or Table 2 or a fragment thereof.

1 28. A method of determining the prognosis of an individual with colorectal  
2 cancer comprising:

3 a) determining the expression of one or more genes selected from the group  
4 consisting of a nucleic acid of Table 1 or Table 2 or a fragment thereof in a first tissue type of  
5 a first individual; and

6 b) comparing said expression of said gene(s) from a second normal tissue type  
7 from said first individual or a second unaffected individual;

8                    wherein a substantial difference in said expression indicates a poor prognosis.

1                    29. A method of treating colorectal cancer comprising administering to an  
2 individual having colorectal cancer an antibody to a colorectal cancer modulator protein or  
3 fragment thereof conjugated to a therapeutic moiety.

1                    30. The method of Claim 29, wherein said therapeutic moiety is a cytotoxic  
2 agent.

1                    31. The method of Claim 29, wherein said therapeutic moiety is a  
2 radioisotope.